Multiple plant-wax compounds record differential sources and ecosystem structure in large river catchments

Jordon D. Hemingway[a,b,*], Enno Schefuß[c], Bienvenu Jean Dinga[d], Helena Pryer[a], Valier V. Galy[a]

[a]Department of Marine Chemistry and Geochemistry, Woods Hole Oceanographic Institution, 266 Woods Hole Road, Woods Hole, MA 02543, USA

[b]Massachusetts Institute of Technology-Woods Hole Oceanographic Institution Joint Program in Oceanography and Applied Ocean Science and Engineering, 77 Massachusetts Avenue, Cambridge, MA 02139, USA

[c]MARUM - Center for Marine Environmental Sciences, Universität Bremen, 28334 Bremen, Germany

[d]Groupe de Recherche en Sciences Exactes et Naturelles / Délégation Générale à la Recherche Scientifique et Technologique (GRSEN/DGRST), Ministere de la Recherche Scientifique, Brazzaville, Congo

*Corresponding Author. Tel.: +1 508 289 2821. *E-mail address*: jhemingway@whoi.edu (J.D. Hemingway)

Abstract

The concentrations, distributions, and stable carbon isotopes ($\delta^{13}C$) of plant waxes carried by fluvial suspended sediments contain valuable information about terrestrial ecosystem characteristics. To properly interpret past changes recorded in sedimentary archives it is crucial to understand the sources and variability of exported plant waxes in modern systems on seasonal to inter-annual timescales. To determine such variability, we present concentrations and $\delta^{13}C$ compositions of three compound classes (*n*-alkanes, *n*-alcohols, *n*-alkanoic acids) in a 34-month time series of suspended sediments from the outflow of the Congo River.

We show that exported plant-dominated *n*-alkanes ($C_{25} – C_{35}$) represent a mixture of C_3 and C_4 end members, each with distinct molecular distributions, as evidenced by an 8.1 ± 0.7‰ (±1σ standard deviation) spread in $\delta^{13}C$ values across chain-lengths, and weak correlations between individual homologue concentrations (r = 0.52 – 0.94). In contrast, plant-dominated *n*-alcohols ($C_{26} – C_{36}$) and *n*-alkanoic acids ($C_{26} – C_{36}$) exhibit stronger positive correlations (r = 0.70 – 0.99) between homologue concentrations and depleted $\delta^{13}C$ values (individual homologues average ≤ -31.3‰ and -30.8‰, respectively), with lower $\delta^{13}C$ variability across chain-lengths (2.6 ± 0.6‰ and 2.0 ± 1.1‰, respectively). All individual plant-wax lipids show little temporal $\delta^{13}C$ variability throughout the time-series (1σ ≤ 0.9‰), indicating that their stable carbon isotopes are not a sensitive tracer for temporal changes in plant-wax source in the Congo basin on seasonal to inter-annual timescales.

Carbon-normalized concentrations and relative abundances of *n*-alcohols (19 – 58% of total plant-wax lipids) and *n*-alkanoic acids (26 – 76%) respond rapidly to seasonal changes in runoff, indicating that they are mostly derived from a recently entrained local source. In contrast, a lack of correlation with discharge and low, stable relative abundances (5 – 16%) indicate that *n*-alkanes better represent a catchment-integrated signal with minimal response to discharge seasonality. Comparison to published data on other large watersheds indicates that this phenomenon is not limited to the Congo River, and that analysis of multiple plant-wax lipid classes and chain lengths can be used to better resolve local vs. distal ecosystem structure in river catchments.

1. Introduction

Since their discovery (Eglinton et al., 1962; Eglinton and Hamilton, 1967), the information recorded in the composition of aliphatic plant-wax lipids has been utilized extensively as a recorder of terrestrial ecosystem structure both in modern settings (Diefendorf et al., 2011; Bush and McInerney, 2013) and the geologic past (see Pancost and Boot, 2004; Eglinton and Eglinton, 2008; Freeman and Pancost, 2014 for review). Much attention has been focused on long-chain (i.e. greater than ~23 carbons) saturated *n*-alkanes, such that the detection of distinct homologue distributions among plant functional types (PFTs) has lead to the use of homologue ratios as a tracer for *n*-alkane sources and ecosystem composition (Ficken et al., 2000; Pancost et al., 2002; Bingham et al., 2010). Such ratios have been frequently utilized in geologic records to infer past ecosystem changes, assuming a straightforward relationship between *n*-alkane production and PFT coverage. However, it has recently been recognized that mixing of *n*-alkanes is likely nonlinear with respect to ecosystem composition, as the absolute production rate of these compounds varies greatly by PFT and between individual species within the same PFT (Rommerskirchen et al., 2006; Vogts et al., 2009; Diefendorf et al., 2011; Magill et al., 2013a; Bush and McInerney, 2013; Garcin et al., 2014). To circumvent these issues, the simultaneous measurement of additional *n*-alkyl lipid classes (i.e. *n*-alcohols and *n*-alkanoic acids) should provide complementary information on plant-wax, and thus terrestrial organic carbon, sources and variability (e.g. Chikaraishi and Naraoka, 2006; Jansen et al., 2006; Diefendorf et al., 2011; Galy et al., 2011; Tao et al., 2015).

Gas chromatography coupled to isotope ratio mass spectrometry (GC-IRMS) allows for the stable carbon isotope ($\delta^{13}C$) analysis of individual compounds (Hayes et al., 1989; Hayes, 1993). Due to their differential fractionation of ^{13}C during photosynthesis, such measurements enable the determination of relative contributions by C_3, C_4 and crassulacean acid metabolism photosynthetic pathways to individual lipids (Collister et al., 1994; Hobbie and Werner, 2004 and references therein). However, it has been shown that competing factors such as light and water stress can cause secondary fractionation effects (e.g. Graham et al., 2014), potentially complicating interpretation of $\delta^{13}C$ compositions and changes thereof.

Combining $\delta^{13}C$ and distribution data, therefore, provides an additional constraint on the mixing of plant-wax lipid sources in environmental samples. For example, $\delta^{13}C$ differences

between homologous lipids of the same compound class as high as ~6‰ have been observed in fluvial sediments due to increasing influence of C_4 grasses at longer chain lengths (Freeman and Colarusso, 2001; Galy et al., 2011; Hötzel et al., 2013; Wang et al., 2013a; Agrawal et al., 2014). In contrast, differences in ^{13}C fractionation between *n*-alkyl lipid classes from the same species have been shown to be negligible (≤1‰) compared to differences between photosynthetic pathways (~13‰; Rommerskirchen et al., 2006; Chikaraishi and Naraoka, 2007; Vogts et al., 2009). Therefore, in addition to their distributions, $\delta^{13}C$ values of multiple lipid classes should act as a more robust constraint on the sources of plant organic matter in environmental samples (e.g. Chikaraishi and Naraoka, 2006; Diefendorf et al., 2011; Galy et al., 2011; Feng et al., 2013; Tao et al., 2015).

Because of their specificity as a plant biomarker, long-chain *n*-alkyl lipids are ideally suited for reconstructing ecosystem changes recorded in terrestrially dominated lacustrine and marine sediments (Pancost and Boot, 2004; Eglinton and Eglinton, 2008; Castañeda and Schouten, 2011; Freeman and Pancost, 2014). For example, *n*-alkyl lipid $\delta^{13}C$ measurements have been used for reconstructions such as savannah land cover response to climate change during the last deglaciation (Hughen et al., 2004) and the Miocene C_4 grassland expansion (Freeman and Colarusso, 2001; Hötzel et al., 2013). However, interpretation of individual compound $\delta^{13}C$ values as a reconstruction of ecosystem land cover is likely complicated by effects such as a nonlinear response to C_3/C_4 coverage (Garcin et al., 2014) and insensitivity to changes within C_3 photosynthetic ecosystems (i.e. woody vs. non-woody; Feakins et al., 2013; Magill et al., 2013a; Magill et al., 2013b).

Additionally, spatial integration is likely not uniform within a river catchment, as changes in plant-wax distribution and isotope signals have been observed during fluvial transit (Galy and Eglinton, 2011; Galy et al., 2011; Ponton et al., 2014). Such non-uniform spatial integration should affect each compound differentially, and will lead to biased reconstructions of catchment land cover depending on which compound is used (e.g. Wang et. al, 2013a). In order to properly interpret paleo-environmental plant-wax signals recorded in sedimentary archives it is therefore crucial to better understand how well various classes of fluvially exported *n*-alkyl lipids represent catchment-integrated vegetation coverage, and on what timescales.

2. Background

The Congo River provides an ideal opportunity to address this question. Draining 3.6 × 10^6 km^2 of central Africa between 15°S - 10°N, the Congo is highly influenced by seasonal monsoonal precipitation due to the north-to-south migration of the inter-tropical convergence zone (ITCZ; Gasse, 2000). Catchment land cover is dominated by nearly equal amounts of closed-canopy evergreen rainforest (31%) and deciduous woodland/shrubland (26%), with lesser amounts of deciduous/montane forest (20%), mixed savannah/grassland (15%) and permanently inundated swamp forest (4%; Mayaux et al., 2004; Still and Powell, 2010). In general, land cover shifts from deciduous woodland/shrubland and mixed savannah/grassland in the headwaters to predominantly evergreen rainforest downstream, although small regions containing woodland/shrubland and savannah/grassland are present near the sampling site (Figure 1a). This corresponds to a shift from a mixed C_3/C_4 signal in both northern and southern hemisphere headwaters to nearly C_3-exclusive land cover near the equator, especially in the main-stem swamp forest (*Cuvette Congolaise*) and its tributaries (Figure 1b).

Congo River discharge (Q_w) is remarkably stable throughout the year due to a seasonal offset in peak northern- and southern-hemisphere contribution, leading to an annual maximum discharge at Brazzaville/Kinshasa equal to roughly double the annual minimum (Coynel et al., 2005; Spencer et al., 2014). High rainfall in the north of the catchment between May and September and a ~1-2 month transit time corresponds to peak discharge of right-bank tributaries during boreal autumn – i.e. September through November (Bricquet, 1993; Mahe, 1993). Combined with increased flow through the *Cuvette Congolaise*, this leads to the observed annual discharge maximum in December (Figure 2a; Bricquet, 1993). In contrast, peak southern-hemisphere rainfall from November through March increases left-bank tributary discharge and is the source of the secondary discharge maximum observed at Brazzaville/Kinshasa (Figure 2a; Bricquet, 1993; Mahe, 1993), thus leading to the increased southern-hemisphere contribution from February through May (Figure 2b; Bricquet, 1993).

This unique spatial separation of PFTs (Figure 1) and temporal separation of tributary discharge (Figure 2b) should lead to pronounced seasonal variability in exported *n*-alkyl lipid source. Here, we aim to address the following questions regarding *n*-alkyl lipids exported in Congo River suspended sediments:

(i.) How do exported lipid signals respond to changes in environmental conditions (i.e. discharge) on seasonal to inter-annual timescales?

(ii.) Are certain lipid classes more representative of specific source regions, and how do lipid classes integrate local vs. distal sources?

(iii.) How can complementary information obtained from multiple compound classes be used to better reconstruct catchment ecosystem coverage and interpret paleo-environmental records?

To do so, we utilize a 34-month time-series of suspended sediments collected near Kinshasa/Brazzaville between November 2010 and August 2013. We combine *n*-alkane, *n*-alcohol, and *n*-alkanoic acid concentrations, distributions, and $\delta^{13}C$ values with simultaneous measurements of total suspended sediment (TSS) concentration, %OC, and river discharge to discern seasonal changes in the source of exported plant waxes.

3. Materials and Methods

3.1. Sample collection

Suspended sediment samples were collected once per month from November 2010 through August 2013 near Brazzaville/Kinshasa, just downstream of Pool Malebo and ~300km upstream of the Congo Estuary (4.3°S, 15.3°E; Figure 1). The sampling location is downstream of all major tributaries, capturing >95% of the total Congo River catchment, and the effect of the downstream Congo Rapids on bulk organic geochemical properties has been shown to be minimal (Spencer et al., 2012). Samples are therefore taken to be representative of material exported to the estuary.

A known volume of surface water (~25L) collected near the center of the channel was filtered through a polyethersulfone (PES) membrane filter (Millipore Corporation) with a nominal pore size of 0.22μm. Filters were dried at 60°C for storage and shipment, and sediments were quantitatively re-suspended in 18MΩ MilliQ water and freeze-dried in pre-combusted glass jars at -40°C (Christ Alpha-L1 equipped with an in-line cold trap) and weighed for TSS concentration. Discharge was measured concurrently with sample collection via a gauging station operated by the Groupe de Recherche en Sciences Exactes et Naturelles (Republic of

Congo) and a rating curve which is periodically updated by Acoustic Doppler Current Profiler (ADCP) transects (Figure 2a). Triplicate transects indicate that precision of discharge measurements is ± 5%, although overbank flooding during periods of high discharge likely biases measurements towards an underestimate of the true value (Spencer et al., 2014).

3.2. Extraction, separation, and purification of n-alkyl lipids

After weighing, sediments were homogenized using an agate mortar and pestle and an aliquot was removed for bulk analysis. One sample (June 2013) contained coarse vegetation debris, which was manually removed using solvent-cleaned forceps and weighed separately (16 mg; not included in extraction). Sediments were then extracted at 100°C for 20 minutes in a microwave accelerated reaction system (MARS, CEM Corporation) in 20mL of dichloromethane (DCM) and methanol (9:1). Total lipid extracts were saponified at 70°C for 2 hours using 0.5M KOH in methanol, after the addition of ~1% 18MΩ MilliQ water to prevent methylation of carboxylic acid functional groups. After the addition of 15mL of water and ~1g pre-combusted NaCl (to increase density difference), "base" fractions were liquid-liquid extracted in 5mL of pure hexane 5 times. Hydrochloric acid was then added until reaching pH 2, and "acid" fractions were extracted using hexane and DCM (4:1) until the organic phase was clear (generally 5 times). Both acid and base fractions were further purified by column chromatography using 1g of Supelclean amino-propyl silica gel (Supelco Analytical) and the following elution scheme: 4mL hexane (F1); 7mL hexane and DCM (4:1, F2); 10mL DCM and acetone (9:1, F3); 14mL 2% (w/w) formic acid in DCM (F4); 18mL DCM and methanol (1:1, F5). Acid and base fractions containing alkanes (F1), alcohols (F3), and alkanoic acids (F4) were recombined to ensure maximum recovery.

To isolate *n*-alkanes, $F1_T$ (acid and base fractions recombined in 1.5mL 2:1 hexane:DCM) was subjected to urea adduction in which 500µL of urea-saturated methanol was added and solvent was evaporated using a stream of N_2 gas to promote urea recrystallization (repeated three times). Crystals were rinsed three times with pure hexane to remove the "non adducted" fraction before being dissolved in 15mL MilliQ water and liquid-liquid extracted using pure hexane as described above. Both alcohols and alkanoic acids require derivatization in order to be amenable to gas chromatography. Alcohols were acetylated in 250µL of pyridine and acetic anhydride with known isotopic composition (1:1) at 70°C for 1 hour. Alkanoic acids were

trans-esterified in 15mL of HCl and methanol with known isotopic composition (5:95) at 70°C for 12 hours. MilliQ water (15mL) was then added, and fatty acid methyl esters (FAMEs) were liquid-liquid extracted into hexane and DCM (4:1) five times. FAMEs were further purified by column chromatography using 1g of amino-propyl silica gel eluted with: 4mL hexane ($F4_TF1$); 7mL hexane and DCM (4:1, $F4_TF2$); 18mL DCM and methanol (1:1, $F4_TF3$).

After quantification but before isotope analysis, unsaturated compounds were removed using 0.5g silver nitrate silica gel (Supelco Analytical) in a Pasteur pipette column eluted with: 5mL hexane and DCM (95:5, SN1); 18mL hexane and DCM (4:1, SN2); 5mL DCM and acetone (1:1, SN3). Fractions containing *n*-alkanes ($F1_T$, adducted), *n*-alcohols ($F3_T$, SN2), and *n*-alkanoic acids ($F4_TF2$, SN2) were stored at 4°C until analyzed. Recovery using this protocol is ~90%, as determined by periodically subjecting a known mixture of compounds to the entire procedure.

3.3. Quantification and isotopic measurements

Total organic carbon (%OC) of bulk sediments was measured after decarbonation over HCl fumes at 60°C for 72 hours using a Fisons elemental analyzer coupled to a Finnigan Delta plus IRMS as described in Whiteside et al., 2011. All *n*-alkyl lipids were quantified using a Hewlett Packard 5890 gas chromatograph-flame ionization detector (GC-FID) with a Gerstel PTV injection system, and separated with a VF-1MS capillary column (Agilent Technologies). Temperature program was as follows: ramp to 130°C at 30°C/min, ramp to 320°C at 8°C/min, hold for 7.5min at 320°C (35min total). Samples were analyzed as a single injection and compared to an external standard run at 3 dilutions between every ~5 samples. Uncertainty was calculated using the standard deviation of the best-fit line to the calibration curve.

Compound-specific $\delta^{13}C$ was determined using a ThermoFisher Scientific Trace GC Ultra with a DB-1MS capillary column (Agilent Technologies) coupled to a Finnigan MAT252 IRMS via a GC/C combustion interface modified for oxygen trickle flow (Merritt et al., 1995; Sessions, 2006). Temperature program was as follows: hold for 3min at 120°C, ramp to 200°C at 30°C/min, ramp to 320°C at 4°C/min, hold for 29.3min at 320°C (70min total). All samples were measured at least in duplicate (triplicate when not limited by low concentrations) and calibrated against pulses of CO_2 gas with a known $\delta^{13}C$ value. Long-term precision of an external *n*-alkane standard mixture was ≤0.2‰ (±1σ standard deviation). Results for individual compounds after

correction for derivatization agent are reported with uncertainty as ±1σ of all injections. Data are reported relative to Vienna Pee-Dee Belemnite (VPDB).

3.4. Data analysis

One sample (September 2013) was omitted from the dataset due to contamination by the PES membrane filter, inhibiting the ability to measure bulk %OC. Additionally, one sample (February 2011) returned spurious $\delta^{13}C$ and concentration values, likely due to improper sampling or influence of a local extreme runoff event, and was thus removed in accordance with Chavenet's criterion (Glover et al., 2011). Regressions were performed as weighted least squares (WLS) using a weighting factor of $1/\sigma_i$ for each sample *i* (Glover et al., 2011). All data analysis was performed in the Python programming language v.2.7 and ArcGIS for desktop v.10.3.

4. Results

4.1. Environmental parameters

All environmental parameters are presented in Table EA1. Congo River discharge recorded at Brazzaville/Kinshasa during the sampling period ranged from a minimum of $23.2\times10^3 \pm 1.1\times10^3$ m^3/s in July 2011 to a maximum of $54.6\times10^3 \pm 2.7\times10^3$ m^3/s in December 2011 (Figure 2a). Average discharge for 2011 (35.3×10^3 m^3/s) was the fifth-lowest since recording began in 1977, while 2012 and 2013 were closer to long-term average values (Spencer et al., 2012). Seasonally, discharge displays two maxima: a large peak in Nov-Dec-Jan during high flow through northern hemisphere tributaries and the main-stem *Cuvette Congolaise*, and a smaller peak in Mar-Apr-May due to increased flow from southern hemisphere tributaries (Bricquet, 1993; Coynel et al., 2005; Bouillon et al., 2012; Spencer et al., 2012; Spencer et al., 2014). This leads to an estimated range in southern-hemisphere contribution (termed f_{south}) of 31 – 53%, with a median value of 39% (Bricquet, 1993). We classify periods with f_{south} above the median value – i.e. February through May – as being "southern hemisphere dominated," and all other times as being "main-stem dominated" or "*Cuvette Congolaise* dominated" (Figure 2b). Importantly, the largest southern-hemisphere tributary, the Kasai River, enters the main-stem downstream of the *Cuvette Congolaise* swamp forest.

TSS concentration averaged 21.1 ± 7.6 g/m^3 throughout the time series, ranging from a

minimum of 10.2 g/m^3 to a maximum of 43.6 g/m^3 (Figure 2c). TSS are rich in carbon, with an average %OC of 6.1 ± 1.0%, leading to an average particulate organic carbon (POC) concentration of 1.3 ± 0.4 g/m^3 with a range of 0.6 – 2.6 g/m^3 (Figure 2d). TSS, POC, and %OC ranges reported here agree well with published values, both for the main-stem Congo as well as the Oubangui River at Bangui Station (Figure 1), a major right-bank tributary (Coynel et al., 2005; Bouillon et al., 2012; Bouillon et al., 2014). Congo River suspended sediment %OC increases slightly as a function of discharge (R^2 = 0.19, p-value = 0.01; not shown), although POC concentration shows no correlation with discharge (p-value = 0.46; not shown).

4.2. Lipid abundance and distribution

Concentrations of individual homologues, average chain lengths, and carbon preference indices are presented in Tables EA2-EA4.

4.2.1. n-Alkanes

Carbon-normalized concentrations of individual plant-wax *n*-alkanes (C_{23} – C_{35}; odd-numbered homologues) range from a minimum of 3.7 ± 0.8 μg/gOC (C_{23}) to a maximum of 82.1 ± 1.3 μg/gOC (C_{29}; Figure 3a). The sum of the long-chain odd-numbered homologue concentrations ($\Sigma LC_{25\text{-}35}$) exhibits considerably less variability, ranging from 66.0 – 207.1 μg/gOC. Time-series plots of $\Sigma LC_{25\text{-}35}$ and selected homologue concentrations are presented in Figure 4a-c. $\Sigma LC_{25\text{-}35}$ concentrations show a slight decrease with increasing discharge (Figure 5a), although this relationship is driven by changes in %OC and disappears when considering sediment-normalized concentrations (R^2 = 0.001, p-value > 0.05; not shown).

n-Alkanes are consistently dominated by C_{29} and C_{31} homologues, contributing up to 33% and 26% to $\Sigma LC_{25\text{-}35}$, respectively. At only 8-9% each, C_{25} and C_{35} are the least abundant homologues, while C_{27} and C_{33} contribute 12-13% each. To compare changes in distributions between samples, we compute the average chain length (ACL) as the concentration-weighted average of C_{25} – C_{35} odd-numbered homologues:

$$\mathrm{ACL} = \frac{25\times[C_{25}]+ 27\times[C_{27}]+ 29\times[C_{29}]+ 31\times[C_{31}]+ 33\times[C_{33}]+ 35\times[C_{35}]}{\Sigma LC_{25-35}} \qquad (1)$$

n-Alkane ACL in our sample set is remarkably stable, with an average of 30.0 ± 0.1 units and a range of 29.6 – 30.2 units, and shows no significant correlation with discharge (Figure 5b). We compute the carbon preference index (CPI) for C_{25} – C_{35}, defined as:

$$\mathrm{CPI} = \frac{1}{2}\left(\frac{\Sigma \mathrm{LC}_{25-35}}{\Sigma \mathrm{LC}_{24-34}} + \frac{\Sigma \mathrm{LC}_{25-35}}{\Sigma \mathrm{LC}_{26-36}}\right) \qquad (2)$$

$\Sigma LC_{25\text{-}35}$ refers to odd-numbered homologues only while $\Sigma LC_{24\text{-}34}$ and $\Sigma LC_{26\text{-}36}$ refer to even-numbered homologues only (we note that C_{36} was not detected in any sample). *n*-Alkane CPI in our dataset averages 2.9 ± 0.5, ranging from 2.1 – 4.1, and shows a small yet statistically significant decrease with increasing discharge (Figure 5c).

We additionally calculate P_{aq}, an estimate of macrophyte contribution to *n*-alkanes (Ficken et al., 2000), as:

$$P_{aq} = \frac{[C_{23}] + [C_{25}]}{[C_{23}] + [C_{25}] + [C_{29}] + [C_{31}]} \qquad (3)$$

Resulting P_{aq} values in our sample set (Table EA2) average 0.19 ± 0.04 with a range of 0.12 – 0.26, and are uncorrelated with discharge (R^2 = 0.08, p-value > 0.05; not shown).

4.2.2. n-Alcohols

While nominally regarded as a plant-wax lipid, C_{24} *n*-alcohol has been observed in freshwater phytoplankton (Volkman et al., 1998; Volkman et al., 1999; Xu et al., 2007). In our sample set, isotopic evidence indicates that phytoplankton contribute to C_{24} *n*-alcohol (see section 5.3. below), and we therefore omit this compound from our calculations of ACL, CPI, and ΣLC.

Plant-wax *n*-alcohols (C_{26} – C_{36}; even-numbered homologues) are considerably more abundant than *n*-alkanes, with individual compound concentrations ranging from 14.6 ± 3.1 μg/gOC (C_{36}) to 163.0 ± 8.0 μg/gOC (C_{28}; Figure 3b). $\Sigma LC_{26\text{-}36}$ concentrations range from 206.5 – 718.7 μg/gOC, and are therefore 3.8 ± 0.9 times higher than corresponding *n*-alkane concentrations. Time series plots of $\Sigma LC_{26\text{-}36}$ and selected homologue concentrations are shown in Figure 6a-c, while Figure 5d shows that $\Sigma LC_{26\text{-}36}$ concentrations decrease as a function of river discharge. Again, this relationship is driven by changes in %OC, as sediment-normalized $\Sigma LC_{26\text{-}36}$ concentrations display no significant relationship with discharge (R^2 = 0.04, p-value > 0.05; not shown).

n-Alcohols are more evenly distributed than *n*-alkanes, with no single homologue contributing more than 21% or less than 9% of the long-chain total (Figure 3b). ACL is calculated similarly to *n*-alkanes, but using C_{26} – C_{36} even-numbered homologues. Again, ACL shows little variability, with a range of 29.8 – 30.6 units and an average of 30.2 ± 0.2 units. CPI,

calculated as above but using $\Sigma LC_{26\text{-}36}$ in the numerator and $\Sigma LC_{25\text{-}35}/\Sigma LC_{27\text{-}37}$ in the denominators (noting that C_{37} was not detected in any sample), averages 3.7 ± 0.4 with a range of 3.0 – 4.6. While ACL shows no correlation (Figure 5e), CPI exhibits a strong negative relationship with discharge (Figure 5f).

4.2.3. n-Alkanoic acids

As C_{24} *n*-alcohol was omitted from the above calculations, to accurately compare ACL, CPI, and ΣLC across compound classes we remove C_{24} *n*-alkanoic acid from the calculations performed here.

Plant-wax *n*-alkanoic acid concentrations (C_{26} – C_{36}; even-numbered homologues) display the highest values and largest variability of all *n*-alkyl lipid classes (Figure 3c). Individual compounds range from 2.7 ± 0.2 μg/gOC (C_{36}) to 457.1 ± 2.3 μg/gOC (C_{28}), with a $\Sigma LC_{26\text{-}36}$ concentration range of 190.2 – 1648.6 μg/gOC. Long-chain *n*-alkanoic acids therefore contribute up to ~0.2% of total exported POC, and are 7.1 ± 2.5 times more abundant than *n*-alkanes. Similar to *n*-alkanes and *n*-alcohols, *n*-alkanoic acid carbon-normalized $\Sigma LC_{26\text{-}36}$ concentrations decrease with increasing river discharge (Figure 5g). While this relationship is partially driven by changes in %OC, sediment-normalized values additionally exhibit a statistically significant decrease ($R^2 = 0.19$, p-value = 1.4×10^{-2}; not shown). Time series plots of selected homologues and $\Sigma LC_{26\text{-}36}$ concentrations are plotted in Figure 7a-c.

n-Alkanoic acids display a similar distribution to *n*-alcohols, with C_{26}, C_{28}, and C_{30} all contributing ~20 – 25% of the long-chain total, and decreasing contribution with increasing chain length beyond C_{30} (Figure 3c). Average ACL is 29.5 ± 0.2, slightly lower than that of *n*-alkanes and *n*-alcohols, and exhibits a slight increase with increasing discharge (Figure 5h). *n*-Alkanoic acid CPI is the highest of all observed compound classes (C_{37} not detected), averaging 4.3 ± 0.5 and shows no correlation with river discharge (Figure 5i). We note that inclusion of *n*-C_{24} decreases ACL to 28.4 ± 0.2 and exhibits no effect on CPI (not shown).

4.3. Compound-specific $\delta^{13}C$

Individual homologue $\delta^{13}C$ measurements are reported in Tables EA5-EA7.

4.3.1. n-Alkanes

n-Alkanes display the largest $\delta^{13}C$ variability across long-chain homologues of all compound classes studied, with an average max-min value of 8.1 ± 0.7‰ (Figure 3d). However, we note that C_{25} could not be measured in two samples (December 2010, July 2013) and C_{35} could not be measured in one sample (July 2013), as concentrations were too low. All samples show the same general trend with chain length – i.e. C_{25}, C_{27}, and C_{33} near -30‰, C_{29} and C_{31} near -34‰, and C_{35} up to -24.7 ± 0.1‰ (Figure 3d). Temporal variability for each compound in the dataset is ~2.5 – 3.0‰ (max-min), as is shown for C_{29} and C_{35} in Figure 4d. $\delta^{13}C$ values of all compounds are uncorrelated with discharge (p-value > 0.05; not shown).

4.3.2. n-Alcohols

Low concentrations prevented the measurement of $\delta^{13}C$ values for C_{36} *n*-alcohol. Additionally, one sample (December 2010) displayed contamination by siloxanes and was omitted. All remaining samples follow the same general pattern, with C_{24} exhibiting the most depleted values, nearly identical values for C_{26} – C_{30} and C_{34}, and C_{32} showing the most enrichment, averaging -31.1 ± 0.7‰ (Figure 3e).

Time series plots of $\delta^{13}C$ values for C_{24} and C_{28} *n*-alcohols are plotted in Figure 6d. C_{24} $\delta^{13}C$ values display a strong positive correlation with discharge (Figure 8a), with the most ^{13}C-depleted value (-36.9 ± 0.1‰) observed during the lowest measured discharge on record (July 2011). In contrast, C_{28} – C_{34} $\delta^{13}C$ values display no correlation with discharge (p-value > 0.05; not shown), although C_{26} exhibits a slight positive relationship, mainly driven by three outlier points (Figure 8b). Resulting $\delta^{13}C$ spread across measured plant-wax *n*-alcohols (i.e. C_{26} – C_{34}) is therefore 2.6 ± 0.6‰, significantly lower than that for *n*-alkanes, even when only considering analogous homologues (i.e. C_{25} – C_{33} *n*-alkane spread of 4.2 ± 0.6‰). While temporal variability within C_{26} – C_{30} homologues is ~2.5‰, C_{32} and C_{34} are significantly more variable, with a max-min value of ~3.5‰.

4.3.3. n-Alkanoic acids

C_{36} *n*-alkanoic acid $\delta^{13}C$ values could not be measured as concentrations were too low, nor could C_{34} in one sample (December 2011). Similar to *n*-alcohols, *n*-alkanoic acids show significantly less spread in $\delta^{13}C$ values between measured homologues (i.e. C_{24} – C_{34}; 2.0 ± 1.1‰) than do *n*-alkanes (Figure 3f). C_{24} and C_{26} *n*-alkanoic acids display the largest temporal

variability of all measured compounds – with a range of 5.2‰ and 5.5‰, respectively. In contrast, C_{28} – C_{32} temporal variability is ~2.5‰, similar to that for *n*-alkanes and *n*-alcohols, while C_{34} varies by 3.4‰ (Figure 7d).

Unlike *n*-alcohols, C_{24} *n*-alkanoic acids do not show ^{13}C-depletion relative to longer chain homologues during periods of low discharge (p-value > 0.05; not shown). In addition, C_{24} – C_{26} *n*-alkanoic acids are ^{13}C-eriched relative to C_{24} *n*-alcohol by 3.9 ± 1.1‰ and 4.3 ± 1.3‰, respectively. C_{24} and C_{28} $\delta^{13}C$ values show a small yet statistically significant enrichment with increasing discharge (Figure 8c, 8d), while all other compounds are uncorrelated (p-value > 0.05; not shown).

4.4. Correlations between homologues and compound classes

WLS regression correlation coefficients (r) and significance p-values for concentrations and $\delta^{13}C$ values of each compound are presented in Tables 1-3.

Within each *n*-alkyl compound class, concentrations of all long-chain homologues exhibit statistically significant positive correlations, with r ranging from of 0.52 – 0.94 for *n*-alkanes, 0.71 – 0.98 for *n*-alcohols, and 0.70 – 0.99 for *n*-alkanoic acids. In contrast, concentrations of long-chain homologues between different compound classes are uncorrelated or display weak positive correlation ($r \leq 0.75$). Concentrations of both C_{23} and C_{25} *n*-alkane are statistically uncorrelated with their corresponding (i.e. n + 1) *n*-alkanoic acids. Additionally, C_{23}, C_{25}, and C_{35} *n*-alkane concentrations are uncorrelated with those of C_{28} – C_{34} *n*-alkanoic acid; C_{36} *n*-alcohol is uncorrelated with C_{30} – C_{36} *n*-alkanoic acid; and C_{29} *n*-alkane is uncorrelated with C_{36} *n*-alcohol, indicating a decoupling between the sources of these compounds.

In general, $\delta^{13}C$ values between all long-chain homologues exhibit less correlation than do concentrations. Within each compound class, r exhibits a range of -0.02 – 0.75 for *n*-alkanes, -0.09 – 0.50 for *n*-alcohols, and -0.55 – 0.79 for *n*-alkanoic acids. Similar ranges are observed between compound classes: -0.09 – 0.78 between *n*-alkanes and *n*-alcohols, -0.16 – 0.63 between *n*-alkanes and *n*-alkanoic acids, and -0.35 – 0.58 between *n*-alcohols and *n*-alkanoic acids. Interestingly, $\delta^{13}C$ values between C_{26} and C_{34} *n*-alkanoic acids display a statistically significant negative correlation, while all other significant correlations are positive.

Lastly, $\delta^{13}C$ values are generally either uncorrelated with concentrations or display a statistically significant negative correlation. C_{27} – C_{33} *n*-alkane, C_{24} *n*-alcohol, and C_{24} *n*-alkanoic

417 acid $\delta^{13}C$ values all exhibit significant negative correlation with increasing concentrations of
418 most measured compounds, while C_{25} *n*-alkane, C_{35} *n*-alkane, C_{28} – C_{30} *n*-alcohol, and C_{34} *n*-
419 alcohol $\delta^{13}C$ values are statistically uncorrelated with the concentrations of all compounds. In
420 contrast to other compound classes, some *n*-alkanoic acid $\delta^{13}C$ values exhibit significant positive
421 correlation with concentrations: C_{30} $\delta^{13}C$ values correlate positively with C_{33} – C_{35} *n*-alkane and
422 C_{36} *n*-alcohol concentrations, C_{32} $\delta^{13}C$ values correlate positively with C_{36} *n*-alcohol
423 concentrations, and C_{34} $\delta^{13}C$ values correlate positively with C_{36} *n*-alkanoic acid concentrations.
424
425 **5. Discussion**
426
427 *5.1. n-Alkane homologues variably record a spatially integrated signal*
428 Contrary to *n*-alcohols and *n*-alkanoic acids, Congo River carbon-normalized *n*-alkane
429 concentrations are relatively low compared to other large rivers studied (van Dongen et al., 2008;
430 Galy et al., 2011; Tao et al., 2015). In addition to vascular plants, petrogenic sources can also
431 contribute to alkanes, especially even chain-length saturated homologues due to the low CPI
432 value (~1.0) of rock-derived sources as compared to plant waxes (Eglinton and Hamilton, 1967;
433 Brooks and Smith, 1969). Low concentrations prevented the measurement of even chain-length
434 $\delta^{13}C$ values in our samples, however CPI values between 2.1 – 4.1 indicate that *n*-alkanes are
435 dominated by a vascular plant signal. Additionally, the Congo catchment is composed mainly of
436 Neoproterozoic craton lithology and exhibits low catchment relief, precluding a significant
437 contribution of outcropped sedimentary rocks to Congo River suspended sediments (Milliman
438 and Farnsworth, 2011; Galy et al., 2015).
439 Due to their lack of functional groups, *n*-alkanes are more resistant to diagenetic
440 degradation within soils and sediments than are *n*-alkanoic acids and *n*-alcohols (Cranwell, 1981;
441 Meyers and Eadie, 1993; Meyers and Ishiwatari, 1993; Sinninghe Damsté et al., 2002; van
442 Dongen et al., 2008). For example, Hoefs et al. (2002) show that *n*-alkanes exhibit ~3× higher
443 preservation factors than do *n*-alkanoic acids upon re-exposure of anoxic sediments to oxygen,
444 while Canuel and Martens (1996) and Sun and Wakeham (1994) calculate lower degradation
445 rates for *n*-alkanes than for functionalized lipids in both oxic and anoxic surface sediments.
446 These results are consistent with observed pre-aging of *n*-alkanes prior to fluvial export, as

447 indicated by ^{14}C-derived ages of plant-wax *n*-alkanes in suspended sediments from other large
448 rivers (e.g. Gustafsson et al., 2011; Tao et al., 2015).
449 Congo River suspended sediment *n*-alkanes exhibit significantly lower CPI values than
450 do *n*-alcohols and *n*-alkanoic acids (Figure 5c, 5f, 5i), suggesting increased exposure to
451 diagenesis (Meyers and Ishiwatari, 1993). Additionally, a compilation of individual African forb,
452 grass, shrub, and tree leaves indicates significant overlap in long-chain (i.e. $\Sigma LC_{25\text{-}35}$, $\Sigma LC_{26\text{-}36}$)
453 plant-wax concentrations and CPI values between compound classes (Table 4). We note that
454 African plant *n*-alkanoic acid concentration measurements are lacking (n = 25; Table 4),
455 potentially leading to the higher mean and median values for this compound class. Inclusion of
456 measurements from shrubs, grasses, and forbs raised in botanical gardens (Gao et al., 2014)
457 lowers this mean value to 406 μg/g dry leaf weight (inter-quartile range of 18 – 643 μg/g dry leaf
458 weight, n = 72), nearly identical to the mean of African plant *n*-alkanes and *n*-alcohols.
459 Therefore, barring extreme biases against the transfer of plant-wax *n*-alkanes into soils and
460 subsequent entrainment into streams, their low and stable relative contribution to total *n*-alkyl
461 lipids in suspended sediments (≤ 16%; Figure 9a, 10a) agrees with relatively stronger exposure
462 to diagenesis as compared to *n*-alcohols and *n*-alkanoic acids.
463 *n*-Alkane concentrations, ACL, CPI, and $\delta^{13}C$ values show little to no correlation with
464 discharge (Figure 4, 5a-5c), indicating that exported *n*-alkane signals do not respond to
465 environmental changes on seasonal timescales. In contrast, if *n*-alkanes were dominated by a
466 recently entrained local signal, discharge should exhibit a strong control on molecular
467 concentration/distribution and/or isotopic composition due to temporal variability in northern vs.
468 southern hemisphere tributary contributions and their corresponding PFT signatures (Figure 1,
469 2b). However, this is not observed. This lack of correlation between *n*-alkane concentration,
470 ACL, and discharge likely explains the similarly low and invariant P_{aq} values (Table EA2), as
471 *Cuvette Congolaise* macrophytes do not contribute significantly to exported *n*-alkanes during
472 periods of high northern hemisphere discharge.
473 Differences in isotopic composition between homologues contain additional information
474 related to residence time and end-member contribution in river systems with stable ecosystems
475 and discharge source regions. Integration over multiple source regions with unique *n*-alkane
476 homologue distributions should result in large $\delta^{13}C$ variability with chain length. For example,
477 Agrawal et al. (2014) observed a consistent increase in $\delta^{13}C$ values with chain length of up to

~6‰ between C_{24} and C_{32} *n*-alkanoic acids in a sediment core taken from the Ganges floodplain at the base of the Himalayas. Additionally, they describe a unique "bimodal" concentration distribution with a maximum at C_{24} and with significantly lower C_{26}/C_{28} and higher C_{30}/C_{32} concentrations than would be expected based on the distributions in modern Ganges suspended sediments (Galy et al., 2011). Taken together, Agrawal et al. (2014) use these results as evidence for degradation of Himalayan C_3 *n*-alkanoic acids and replacement by local floodplain C_4-derived compounds, and conclude that C_{26}/C_{28} better retain a headwater signal while C_{30}/C_{32} exhibit significant overprinting due to higher production of the longer-chain homologues by local C_4 grasses.

In the Congo River, integration of *n*-alkanes over multiple source regions should result in a similarly large $\delta^{13}C$ difference across homologues and lower correlation between homologue concentrations, as is observed (Figure 3d, 4d; Table 1). Depleted $\delta^{13}C$ values for C_{29} and C_{31} *n*-alkane confirm the importance of a C_3 source to these compounds, while relatively ^{13}C-enriched C_{33} and, especially, C_{35} values indicate a larger contribution by C_4 grasses with increasing chain length. This agrees with measurements of individual plant leaves, as African gramminoids have been shown to produce higher relative concentrations of C_{33} (C_{35} not measured) as compared to African trees and forbs (Rommerskirchen et al., 2006; Vogts et al., 2009). Using a typical end-member *n*-alkane isotopic value of -35‰ for C_3 and -22‰ for C_4 plants (e.g. Castañeda and Schouten, 2011), this results in a C_4 contribution as high as 69 ± 6% to C_{35} *n*-alkane and as low as 8 ± 4% to C_{29} *n*-alkane, while remote-sensing results indicate that catchment-wide C_4 gramminoid coverage is ~14% (Figure 1b; Still and Powell, 2010). However, we note that remote sensing likely underestimates C_4 coverage in forested areas, as C_4 plants are masked by C_3 forest canopy.

Spatially, C_4-bearing savannah and woodland/shrubland ecosystems are mostly located at the northern and southern extremes of the catchment, above 5°N and between 5-10°S, while C_3-dominated evergreen forest, deciduous/montane forest, and swamp forest occupy the central region (Figure 1). This geographic separation indicates a variable apparent integration region for *n*-alkane homologues, with the longest chain-length *n*-alkanes biasing toward headwater regions due to higher production by gramminoids. Additionally, observed negative correlations between $C_{27} - C_{33}$ *n*-alkane concentrations and $\delta^{13}C$ values (Table 3) are further evidence for an overprinting of distal C_4 sources during transit. This relationship is strongest for C_{29} and C_{31} (r =

509 -0.74, -0.65), consistent with significant production of these compounds in C_3 trees and forbs
510 (Rommerskirchen et al., 2006; Vogts et al., 2009). In contrast, C_{35} $\delta^{13}C$ values are uncorrelated
511 with concentration, further indicating a predominantly headwater C_4 source to this compound,
512 irrespective of concentration. While regions of mosaic savannah/grassland and deciduous
513 woodland/shrubland exist near the sampling site, especially in left-bank tributaries (Figure 1),
514 this contribution is likely minimal. If local C_4 sources were important, this would lead to ^{13}C-
515 enrichment of all *n*-alkane homologues, especially during periods of predominantly southern
516 hemisphere discharge, which is not observed (Figure 4d).
517 Biasing of C_{33} and C_{35} *n*-alkanes toward a headwater C_4 signal agrees with time series
518 measurements of the Oubangui River at Bangui Station (4°21.8' N, 18°33.1' E; Figure 1).
519 Bouillon et al. (2012) show enriched POC $\delta^{13}C$ values (-26.2 ± 0.4‰, n = 11) during periods of
520 high discharge, when autochthonous production is negligible, as this headwater tributary
521 contains significant amounts of dry woody savannah ecosystem coverage. Additionally, enriched
522 POC $\delta^{13}C$ values up to -22.8‰ have been reported for a small savannah tributary to the
523 Oubangui River, while the nearby savannah-dominated Niari River exhibits POC $\delta^{13}C$ values as
524 high as -18.6‰ (Mariotti et al., 1991; Bouillon et al., 2014). In contrast, the Congo main-stem
525 near Brazzaville displays more depleted POC $\delta^{13}C$ values, averaging -28.2 ± 0.4‰ (n = 5;
526 Spencer et al., 2012).
527 Additional evidence for variable spatial integration of *n*-alkane homologues comes from
528 a positive correlation between the $\delta^{13}C$ values of *n*-alkanoic acids/*n*-alcohols and their
529 corresponding *n*-alkanes (i.e. carbon number - 1), as decarboxylation and dehydration of
530 functionalized *n*-alkyl lipids has been shown to occur rapidly in sediments (Cranwell, 1981; Sun
531 and Wakeham, 1994; Sun et al., 1997; Hoefs et al., 2002). Such relationships are especially
532 strong between C_{30}/C_{32} *n*-alcohol and *n*-alkanoic acid and C_{29}/C_{31} *n*-alkane (r up to 0.75; Table
533 2), indicating that diagenetic contribution by functionalized C_3 plant waxes contributes to the
534 overprinting of these compounds during transit. Taken together, the observed depleted $\delta^{13}C$
535 values, negative correlations between $\delta^{13}C$ and concentration, and positive $\delta^{13}C$ correlations with
536 corresponding functionalized lipids indicate that C_{29} and C_{31} *n*-alkane exhibit significant
537 overprinting during fluvial transit and bias toward a more local signal. In contrast, enriched $\delta^{13}C$
538 values and weaker correlation between $\delta^{13}C$ and concentration for C_{33} and, especially, C_{35} *n*-
539 alkanes are strong evidence that these homologues better retain a headwater signal.

5.2. n-Alcohols and n-alkanoic acids are controlled by recently entrained local sources

Contrary to *n*-alkanes, carbon-normalized concentrations of plant-wax *n*-alcohols and *n*-alkanoic acids in Congo River POC are equal to or greater than the highest observed values in any large river system to date (Saliot et al., 2001; van Dongen et al., 2008; Galy et al., 2011; Tao et al., 2015). Such high *n*-alcohol:*n*-alkane (3.8 ± 0.9) and *n*-alkanoic acid:*n*-alkane (7.1 ± 2.5) ratios in suspended sediments contrast with the overlapping range in concentrations between compound classes found in African plants (Table 4). Additionally, despite an identical range in CPI for *n*-alkanes and *n*-alcohols (no *n*-alkanoic data exist) in individual African plant leaves (Table 4), both functionalized compound classes exhibit higher CPI values than do *n*-alkanes in suspended sediment (Tables EA2-EA4), as diagenetic degradation has been shown to lower CPI (Meyers and Ishiwatari, 1993).

Assuming no pervasive biases against the transfer of *n*-alkanes into soils and subsequent entrainment into streams, high concentrations and CPI values of functionalized lipids relative to *n*-alkanes, despite similar input composition from plants (Table 4), supports the hypothesis that exported *n*-alcohols and *n*-alkanoic acids are mostly sourced from local surface soils with less exposure to diagenesis prior to export. Functionalized wax lipids are known to experience rapid diagenetic dehydration and decarboxylation in sediments (Meyers and Eadie, 1993; Sun and Wakeham, 1994; Canuel and Martens, 1996; Sun et al., 1997). For example, Sun et al. (1997) report that 90% of ^{14}C-labeled C_{16} *n*-alkanoic acids (labeled in the methyl position) are degraded due to decarboxylation within 80 days during incubation experiments. While C_{16} *n*-alkanoic acid is produced ubiquitously in the environment, rapid degradation has additionally been observed for plant-wax-specific *n*-alkanoic acids (i.e. $C_{26} - C_{30}$) and *n*-alcohols ($C_{26} - C_{30}$) upon re-exposure of sediments to oxygen (Hoefs et al., 2002). *n*-Alkanoic acids and *n*-alcohols frequently exhibit the lowest preservation of all lipids in marine and lacustrine sediments and have been observed to degrade at faster rates than bulk OC (Cranwell, 1981; Meyers and Ishiwatari, 1993). However, lipid preservation is additionally a function of sediment mineralogy, as sorptive interactions with mineral surfaces have been shown to stabilize labile OC (e.g. Keil et al., 1994; Mayer 1994).

Isotopic evidence further indicates a predominantly local source, as these compound classes exhibit depleted $\delta^{13}C$ values for all plant-wax homologues (average ≤ -31.3‰ and -

30.8‰, respectively; Figure 3e-f). Using a C_3 end-member value of -35‰ and a C_4 end-member value of -22‰, as above (Castañeda and Schouten, 2011), this leads to a minimum C_3 contribution to *n*-alcohols of 73 ± 5% (C_{32}) and 68 ± 6% to *n*-alkanoic acids (C_{26}). However, this is likely an underestimate, as relatively enriched $\delta^{13}C$ values for individual C_3 angiosperm lipids have been reported (i.e. up to -30‰; Diefendorf et al., 2011; Garcin et al., 2014). Isotopic evidence therefore indicates that functionalized lipids are predominantly sourced from local C_3 ecosystems, as C_4 land cover is mostly limited to distal headwater regions (Figure 1b). Similar to *n*-alkanes, regions of mosaic savannah/grassland and deciduous woodland/shrubland near the sampling site likely do not contribute significantly to exported *n*-alcohols and *n*-alkanoic acids, as this would lead to a ^{13}C-enrichment during southern hemisphere dominated discharge periods, which is not observed (Figure 6d, 7d, 8).

Unlike longer chain homologues, autochthonous production of C_{24} *n*-alcohol has been observed in freshwater phytoplankton (Volkman et al., 1998; Volkman et al., 1999; Xu et al., 2007) and is likely a significant source of this compound in our sample set. This is supported by depleted $\delta^{13}C$ values (Figure 3e) and a strong positive relationship with discharge (Figure 8a). If dissolved inorganic carbon (DIC) is ^{13}C-depleted relative to atmospheric CO_2, autochthonous contribution will lead to lower observed $\delta^{13}C$ values for C_{24} *n*-alcohol, especially during periods of low discharge when phytoplankton production is highest. While no DIC $\delta^{13}C$ values exist at our sampling site, low- and rising-water values at Bangui station average -10.0 ± 2.2‰ (n = 30; Bouillon et al., 2012; Bouillon et al., 2014). Additionally, C_{24} *n*-alcohol $\delta^{13}C$ values are strongly correlated with those of C_{22} *n*-alcohol ($R^2 = 0.75$, p-value = 4.0×10^{-10}; not shown), the dominant lipid in freshwater phytoplankton (Volkman et al., 1998; Volkman et al., 1999; Xu et al., 2007), and are uncorrelated with longer chain-length values (p-value > 0.05; not shown). While a slight $\delta^{13}C$ vs. discharge correlation is observed for other compounds (i.e. C_{26} *n*-alcohol, C_{24} and C_{28} *n*-alkanoic acid; Figure 8), these homologues are consistently ~3-5‰ enriched relative to C_{24} *n*-alcohol, indicating minimal autochthonous contribution.

Further evidence for a local, C_3 signal to functionalized *n*-alkyl lipids comes from the fact that $\delta^{13}C$ values show significantly weaker negative correlation with lipid concentrations than do *n*-alkanes, with the exception of C_{24} *n*-alcohol and C_{24} *n*-alkanoic acid (Table 3). As with *n*-alkanes, a negative correlation would indicate addition of C_3 material to a background C_4 signal during transit. However, this is not the case, especially for longer chain-length homologues (i.e.

C_{28+}), indicating negligible contribution by C_4-dominated headwater ecosystems to measured compounds and therefore a smaller apparent integration region than is observed for *n*-alkanes, especially C_{33} and C_{35}. African C_4 gramminoids exhibit similar *n*-alcohol and *n*-alkanoic acid production rates as African forbs, shrubs, and trees (Ali et al., 2005a; Rommerskirchen et al., 2006; Vogts et al., 2009), indicating that this signal is not due to a source effect. Rather, it is likely the result of quantitative diagenetic degradation of headwater functionalized *n*-alkyl lipids during fluvial transit (Cranwell, 1981; Meyers and Ishiwatari, 1993; Sun et al., 1997; Hoefs et al., 2002; van Dongen et al., 2008). In addition, a low spread in $\delta^{13}C$ values across plant-wax chain-lengths (Figure 3e-f) and strong positive correlations between homologue concentrations (Table 1) precludes significant spatial integration of multiple PFTs with unique molecular distribution and isotope composition (c.f. Agrawal et al., 2014).

Additionally, we observe large seasonal variability in *n*-alcohol and *n*-alkanoic acid relative contribution (Figure 9), indicating a change in functionalized lipid source in response to seasonal hydrology. This is consistent with the above evidence that these compounds are sourced from recently entrained OC and integrate a mostly local signal. *n*-Alcohol relative contribution displays a statistically significant increase during *Cuvette Congolaise* dominated periods, balanced by an equal decrease in *n*-alkanoic acids (Figure 10). These results agree with literature measurements of individual plants, as macrophytes display considerably higher *n*-alcohol production rates than do other PFTs (Ficken et al., 1998; Ficken et al., 2000; Bugalho et al., 2004; Ali et al., 2005a; Ali et al., 2005b; Aichner et al., 2010; Gao et al., 2011; Diefendorf et al., 2011; Wang and Liu, 2012; Gao et al., 2014).

It has been shown that the Congo main-stem and a range of tributaries bias toward swamp-forest-like chemical properties during periods of high discharge, indicating an increased contribution by this ecosystem to exported organic carbon (Wang et al., 2013b; Mann et al., 2014). These observations, combined with an increase in *n*-alcohol fractional contribution, are strong evidence for a significant increase in *Cuvette Congolaise* contribution to functionalized *n*-alkyl lipids during periods of high northern hemisphere discharge. Thus, our results indicate that this geographically small region (4% coverage; Mayaux et al., 2004) exhibits a dominant control on the composition of exported functionalized *n*-alkyl lipids in response to seasonal changes in hydrology. However, a lack of significant $\delta^{13}C$ variability across the time series for any

functionalized plant-wax lipid (Figure 6d, 7d) indicates that their $\delta^{13}C$ values are not a sensitive tracer for changes in *n*-alkyl lipid source on these timescales.

5.3. Comparison to other river basins and global significance

Variable spatiotemporal integration of *n*-alkane homologues and a local, recently entrained *n*-alcohol/*n*-alkanoic acid signal is a feature not limited to the Congo River catchment. For example, isotopic and molecular signals of *n*-alkanes and *n*-alkanoic acids show differential contribution by C_4 grasses during transport in the Ganges River through the floodplain (Galy et al., 2011). Using the data of Galy et al. (2011), we compare the concentration-weighted Himalayan plant-wax signal to that just before the confluence with the Brahmaputra River in Bangladesh, noting that sediment fluxes are nearly identical in all major Himalayan tributaries (Andermann et al., 2012).

Himalayan plant-wax *n*-alkanes ($C_{25} - C_{35}$) and *n*-alkanoic acids ($C_{26} - C_{34}$) display nearly identical $\delta^{13}C$ composition, averaging -32.1‰ and -32.3‰ respectively, with similar spread between chain-lengths of 1.8‰ and 1.4‰. In contrast, downstream Ganges *n*-alkanoic acid $\delta^{13}C$ values are enriched by an average of 1.3‰ relative to *n*-alkanes. Additionally, isotopic spread between chain lengths remains constant for *n*-alkanoic acids (i.e. 1.4‰), but increases to 3.5‰ for *n*-alkanes, while ACL of both compound classes increases by ~1 unit. Combined, these results indicate that C_3 *n*-alkanoic acids sourced in the Himalayan range are quantitatively replaced by a mixed C_3/C_4 floodplain signal independent of chain length. In contrast, *n*-alkanes display differential contribution by a floodplain signal across chain lengths, with C_{33}/C_{35} showing the most influence. Quantitative *n*-alkanoic acid replacement during floodplain transit agrees with the results of Agrawal et al. (2014), which already show significant overprinting near the base of the Himalayan range. Thus, despite large differences in sediment erosion rates and biospheric carbon yields between the Congo and Ganges rivers (Galy et al., 2015), exported plant waxes display similar behavior in these two catchments.

In addition to the Ganges River, a large spread in *n*-alkane $\delta^{13}C$ values with chain-length has been observed in settings such as Cameroonian lacustrine and Washington margin surface sediments (Feng et al., 2013; Garcin et al., 2014) and a Zambezi River sedimentary archive (Wang et al., 2013a). Differential contribution by C_3/C_4 plants to *n*-alkanes with chain length therefore appears to be a common phenomenon. We suggest that $\delta^{13}C$ measurement of multiple

n-alkane chain lengths can be used to address nonlinear PFT mixing during transport (Garcin et al., 2014), as C_{33} and, especially, C_{35} bias almost exclusively toward a C_4 end-member, opposite to C_{29} and C_{31}.

Additionally, differential sourcing between compound classes (i.e. functionalized vs. *n*-alkanes) appears to be common in river catchments spanning multiple ecosystems. Thus, in addition to *n*-alkanes, measurement of *n*-alcohols and *n*-alkanoic acids in river sediments and fluvially dominated sedimentary archives can be utilized to address geospatial PFT distribution within the catchment, as functionalized lipids will bias toward a local signal (e.g. Galy et al., 2011; Ponton et al., 2014; this study).

6. Conclusion

We report concentrations and $\delta^{13}C$ values of three classes of dominantly plant-derived *n*-alkyl lipids from a 34-month time series of Congo River suspended sediments. Our results show that *n*-alkanoic acid and *n*-alcohol concentrations are equal to or greater than the highest OC-normalized concentrations in large fluvial systems reported to date. In contrast, *n*-alkanes concentrations are lower than those reported in other major rivers. Spread in *n*-alcohol and *n*-alkanoic acid $\delta^{13}C$ values between long-chain homologues is lower than observed in other major rivers, while *n*-alkanes exhibit up to ~8‰ enrichment with increasing chain length.

These data indicate that *n*-alkanoic acids and *n*-alcohols are sourced from local, C_3-dominated ecosystems, consistent with the idea that high reactivity of functional groups precludes significant spatial integration of these compounds. In contrast, *n*-alkane homologues variably integrate over a wide range of ecosystems with increasing contribution by distal C_4-dominated savannah and woodland/shrubland source regions to the longest chain-length compounds. Strong seasonal shifts in relative *n*-alkanoic acid and *n*-alcohol concentrations indicate that functionalized lipids respond rapidly to changes in hydrological regime. This signal, however, is not reflected in $\delta^{13}C$ values. During periods of highest northern hemisphere discharge, an increase in fractional *n*-alcohol contribution and decrease in *n*-alkanoic acid contribution suggest a strong bias towards a local swamp-forest signal. *n*-Alkanes are less affected by seasonal changes in discharge, further indicating that these compounds integrate over a larger source region.

693 Consequently, we suggest that simultaneous measurement of multiple *n*-alkyl lipid
694 classes and chain lengths in down-core samples will likely provide better geospatial resolution
695 for paleo-ecosystem reconstruction due to their differential integration regions and C_3/C_4 biases.
696
697 **Acknowledgements**
698 We thank Carl Johnson (WHOI), Sarah Rosengard (WHOI), and Ralph Kreutz
699 (MARUM) for laboratory assistance. J.D.H. was supported by the National Science Foundation
700 Graduate Research Fellowship under Grant No. 2012126152. V.V.G. was partly supported by
701 the US National Science Foundation, grants OCE-0851015 and OCE-0928582. Parts of this work
702 were supported by the DFG Research Center/Cluster of Excellence "The Ocean in the Earth
703 System" at MARUM - Center for Marine Environmental Science, University of Bremen. This
704 manuscript benefited greatly through the constructive comments of Aaron Diefendorf, Clayton
705 Magill, one anonymous reviewer, and associate editor Thomas Bianchi.

706 **References**

707 Agrawal S., Galy V. V., Sanyal P. and Eglinton T. I. (2014) C_4 plant expansion in the Ganga
708 Plain during the last glacial cycle: Insights from isotopic composition of vascular plant
709 biomarkers. *Org. Geochem.* **67**, 58–71.

710 Aichner B., Herzschuh U. and Wilkes H. (2010) Influence of aquatic macrophytes on the stable
711 carbon isotopic signatures of sedimentary organic matter in lakes on the Tibetan Plateau.
712 *Org. Geochem.* **41**, 706–718.

713 Ali H. A. M., Mayes R. W., Hector B. L. and Ørskov E. R. (2005a) Assessment of *n*-alkanes,
714 long-chain fatty alcohols and long-chain fatty acids as diet composition markers: The
715 concentrations of these compounds in rangeland species from Sudan. *Anim. Feed Sci. Tech.*
716 **121**, 257–271.

717 Ali H. A. M., Mayes R. W., Hector B. L., Verma A. K. and Ørskov E. R. (2005b) The possible
718 use of *n*-alkanes, long-chain fatty alcohols and long-chain fatty acids as markers in studies of
719 the botanical composition of the diet of free-ranging herbivores. *J. Agr. Sci.* **143**, 85–95.

720 Andermann C., Crave A., Gloaguen R., Davy P., and Bonnet S. (2012) Connecting source and
721 transport: Suspended sediments in the Nepal Himalayas. *Earth Planet. Sc. Lett.* **351-352**,
722 158-170.

723 Bingham E. M., McClymont E. L., Väliranta M., Mauquoy D., Roberts Z., Chambers F. M.,
724 Pancost R. D. and Evershed R. P. (2010) Conservative composition of *n*-alkane biomarkers
725 in Sphagnum species: Implications for palaeoclimate reconstruction in ombrotrophic peat
726 bogs. *Org. Geochem.* **41**, 214–220.

727 Bouillon S., Yambélé A., Spencer R. G., Gillikin D. P., Hernes P. J., Six J., Merckx R. and
728 Borges A. V. (2012) Organic matter sources, fluxes and greenhouse gas exchange in the
729 Oubangui River (Congo River basin). *Biogeosciences* **9**, 2045–2062.

730 Bouillon S., Yambélé A., Gillikin D. P., Teodoru C., Darchambeau F., Lambert T. and Borges A.
731 V. (2014) Contrasting biogeochemical characteristics of the Oubangui River and tributaries
732 (Congo River basin). *Sci. Rep.* **4**: 5402.

733 Bricquet (1993) Les ecoulements du Congo a Brazzaville et la spatialization des apports. *Grands*
734 *Bassins Fluviaux*, Paris, 27-38.

735 Brooks J. D. and Smith J. W. (1969) The diagenesis of plant lipids during the formation of coal,
736 petroleum and natural gas—II. Coalification and the formation of oil and gas in the

Gippsland Basin. *Geochim. Cosmochim. Ac.* **33**, 1183–1194.

Bugalho M. N., Dove H., Kelman W., Wood J. T. and Mayes R. W. (2004) Plant wax alkanes and alcohols as herbivore diet composition markers. *Rangeland Ecol. Manage.* **57**, 259–268.

Bush R. T. and McInerney F. A. (2013) Leaf wax *n*-alkane distributions in and across modern plants: Implications for paleoecology and chemotaxonomy. *Geochim. Cosmochim. Ac.* **117**, 161–179.

Canuel E. A. and Martens C. S. (1996) Reactivity of recently deposited organic matter: Degradation of lipid compounds near the sediment-water interface. *Geochim. Cosmochim. Ac.* **60**, 1793-1806.

Castañeda I. S. and Schouten S. (2011) A review of molecular organic proxies for examining modern and ancient lacustrine environments. *Quaternary Sci. Rev.* **30**, 2851–2891.

Chikaraishi Y. and Naraoka H. (2006) Carbon and hydrogen isotope variation of plant biomarkers in a plant-soil system. *Chem. Geol.* **231**, 190-202.

Chikaraishi Y. and Naraoka H. (2007) $\delta^{13}C$ and δD relationships among three *n*-alkyl compound classes (*n*-alkanoic acid, *n*-alkane and *n*-alkanol) of terrestrial higher plants. *Org. Geochem.* **38**, 198–215.

Collister J. W., Rieley G., Stern B., Eglinton G. and Fry B. (1994) Compound-specific $\delta^{13}C$ analyses of leaf lipids from plants with differing carbon dioxide metabolisms. *Org. Geochem.* **21**, 619–627.

Coynel A., Meybeck M., Seyler P., Etcheber H. and Orange D. (2005) Spatial and seasonal dynamics of total suspended sediment and organic carbon species in the Congo River. *Global Biogeochem. Cyc.* **19.4,** GB4019.

Cranwell P. A. (1981) Diagenesis of free and bound lipids in terrestrial detritus deposited in a lacustrine sediment. *Org. Geochem.* **3**, 79-89.

Diefendorf A. F., Freeman K. H., Wing S. L. and Graham H. V. (2011) Production of *n*-alkyl lipids in living plants and implications for the geologic past. *Geochim. Cosmochim. Ac.* **75**, 7472–7485.

Eglinton G., Gonzalez A. G., Hamilton R. J., and Raphael R. A. (1962) Hydrocarbon constituents of the wax coatings of plant leaves: A taxonomic survey. *Phytochemistry* **1**, 89-102.

Eglinton G. and Hamilton R. J. (1967) Leaf epicuticular waxes. *Science* **156**, 1322-1335.

768 Eglinton T. I. and Eglinton G. (2008) Molecular proxies for paleoclimatology. *Earth Planet. Sc.*
769 *Lett.* **275**, 1–16.
770 Feng X., Benitez-Nelson B. C., Montluçon D. B., Prahl F. G., McNichol A., Xu L., Repeta D. J.,
771 and Eglinton T. I. (2013) ^{14}C and ^{13}C characteristics of higher plant biomarkers in
772 Washington margin surface sediments. *Geochim. Cosmochim. Ac.* **105**, 14-30.
773 Feakins S. J., Levin N. E., Liddy H. M., Sieracki A., Eglinton T. I., Bonnefille R. (2013)
774 Northeast African vegetation change over 12 m.y. *Geology*. **41**, 295-298.
775 Ficken K. J., Barber K. E. and Eglinton G. (1998) Lipid biomarker, $\delta^{13}C$ and plant macrofossil
776 stratigraphy of a Scottish montane peat bog over the last two millennia. *Org. Geochem.* **28**,
777 217–237.
778 Ficken K. J., Li B., Swain D. L. and Eglinton G. (2000) An *n*-alkane proxy for the sedimentary
779 input of submerged/floating freshwater aquatic macrophytes. *Org. Geoch.* **31**, 745–749.
780 Freeman K. H. and Colarusso L. A. (2001) Molecular and isotopic records of C_4 grassland
781 expansion in the late Miocene. *Geochim. Cosmochim. Ac.* **65**, 1439–1454.
782 Freeman K. H. and Pancost R. D. (2014) Biomarkers for terrestrial plants and climate. In
783 *Treatise on Geochemistry (Second Edition)* (eds. Turekian H. D. and Holland K. K.).
784 Elsevier, Oxford. pp. 395 - 416.
785 Galy V. V. and Eglinton T. I. (2011) Protracted storage of biospheric carbon in the Ganges-
786 Brahmaputra basin. *Nat. Geosci.* **4**, 843–847.
787 Galy V. V., Eglinton T. I., France-Lanord C. and Sylva S. P. (2011) The provenance of
788 vegetation and environmental signatures encoded in vascular plant biomarkers carried by the
789 Ganges-Brahmaputra rivers. *Earth Planet. Sc. Lett.* **304**, 1–12.
790 Galy V. V., Peucker-Ehrenbrink B., and Eglinton T. I. (2015) Global carbon export from the
791 terrestrial biosphere controlled by erosion. *Nature*. **521**, 204-207.
792 Gao L., Hou J., Toney J., MacDonald D. and Huang Y. (2011) Mathematical modeling of the
793 aquatic macrophyte inputs of mid-chain *n*-alkyl lipids to lake sediments: Implications for
794 interpreting compound specific hydrogen isotopic records. *Geochim. Cosmochim. Ac.* **75**,
795 3781–3791.
796 Gao L., Edwards E. J., Zeng Y. and Huang Y. (2014) Major evolutionary trends in hydrogen
797 isotope fractionation of vascular plant leaf waxes. *PLoS ONE* **9**, e112610–10.
798 Garcin Y., Schefuß E., Schwab V. F., Garreta V., Gleixner G., Vincens A., Todou G., Séné O.,

Onana J.-M., Achoundong G. and Sachse D. (2014) Reconstructing C_3 and C_4 vegetation cover using *n*-alkane carbon isotope ratios in recent lake sediments from Cameroon, Western Central Africa. *Geochim. Cosmochim. Ac.* **142**, 482–500.

Gasse F. (2000) Hydrological changes in the African tropics since the Last Glacial Maximum. *Quaternary Sci. Rev.* **19**, 189–211.

Glover D. M., Jenkins W. J. and Doney S. C. (2011) Measurement theory, probability distributions, error propagation and analysis. In *Modeling Methods for Marine Science*. Cambridge University Press, New York. pp. 14-45.

Graham H. V., Patzkowsky M. E., Wing S. L., Parker G. G., Fogel M. L., and Freeman K. H. (2014) Isotopic characteristics of canopies in simulated leaf assemblages. *Geochim. Cosmochim. Ac.* **144**, 82-95.

Gustafsson, Ö., Van Dongen, B. E., Vonk, J. E., Dudarev, O. V., & Semiletov, I. P. (2011). Widespread release of old carbon across the Siberian Arctic echoed by its large rivers. *Biogeosciences*, **8**, 1737-1743.

Hayes J. M., Freeman K. H., Popp B. N. and Hoham C. H. (1989) Compound-specific isotopic analyses: A novel tool for reconstruction of ancient biogeochemical processes. *Org. Geochem.* **16**, 1115–1128.

Hayes J. M. (1993) Factors controlling ^{13}C contents of sedimentary organic compounds: Principles and evidence. *Mar. Geol.* **113**, 111–125.

Hobbie E. A. and Werner R. A. (2004) Intramolecular, compound - specific, and bulk carbon isotope patterns in C_3 and C_4 plants: a review and synthesis. *New Phytol.* **161**, 371–385.

Hoefs M. J. L., Rijpistra W. I. C., and Sinninghe Damsté J. S. (2002) The influence of oxic degradation on the sedimentary biomarker record I: Evidence from Madeira Abyssal Plain turbidites. *Geochim. Cosmochim. Ac.* **66**, 2719-2735.

Hötzel S, Dupont LM, Schefuß E, Rommerskirchen F and Wefer G (2013) The role of fire in Miocene to Pliocene C_4 grassland and ecosystem evolution. *Nature Geoscience* **6**, 1027-1030.

Hughen K, Eglinton T. I., Xu L. and Makou M. (2004) Abrupt tropical vegetation response to rapid climate changes. *Science*. **304**, 1955-1959.

Jansen B., Nierop K. G. J., Hageman J. A., Cleef A. M. and Verstraten J. M. (2006) The straight-chain lipid biomarker composition of plant species responsible for the dominant biomass

production along two altitudinal transects in the Ecuadorian Andes. *Org. Geochem.* **37**, 1514–1536.

Keil, R. G., Montluçon, D. B., Prahl, F. G., and Hedges J. I. (1994). Sorptive preservation of labile organic matter in marine sediments. *Nature*, **370**, 549-552.

Magill C. R., Ashley G. M. and Freeman K. H. (2013a) Ecosystem variability and early human habitats in eastern Africa. *P. Natl. Acad. Sci.* **110**, 1167–1174.

Magill C. R., Ashley G. M. and Freeman K. H. (2013b) Water, plants, and early human habitats in eastern Africa. *P. Natl. Acad. Sci.* **110**, 1175–1180.

Mahe (1993) Modulation annuelle et fluctuation interannuelles des precipitations sur le bassin versant du Congo. *Grands Bassins Fluviaux*, Paris, 13-26.

Mann P. J., Spencer R. G., Dinga B. J., Poulsen J. R., Hernes P. J., Fiske G., Salter M. E., Wang Z. A., Hoering K. A., Six J. and Holmes R. M. (2014) The biogeochemistry of carbon across a gradient of streams and rivers within the Congo Basin. *J. Geophys. Res. Biogeosci.* **119**, 687–702.

Mariotti A., Gadel F., Giresse P. and Kinga-Mouzeo (1991) Carbon isotope composition and geochemistry of particulate organic matter in the Congo River (Central Africa): application to the study of Quaternary sediments off the mouth of the river. *Chem. Geol.* **86**, 345–357.

Mayaux P., Martholomé E., S F. and Belwand A. (2004) A new land-cover map of Africa for the year 2000. *J. Biogeogr.* **31**, 861–877.

Mayer, L. M. (1994). Relationships between mineral surfaces and organic carbon concentrations in soils and sediments. *Chem. Geol.*, **114**, 347-363.

Merritt D. A., Freeman K. H., Ricci M. P., Studley S. A. and Hayes J. M. (1995) Performance and optimization of a combustion interface for isotope ratio monitoring gas chromatography/mass spectrometry. *Anal. Chem.* **67**, 2461–2473.

Meyers P. A. and Eadie B. J. (1993) Sources, degradation and recycling of organic matter associated with sinking particles in Lake Michigan. *Org. Geochem.* **20**, 47–56.

Meyers P. A. and Ishiwatari R. (1993) Lacustrine organic geochemistry--an overview of indicators of organic matter sources and diagenesis in lake sediments. *Org. Geochem.* **20**, 867-900.

Milliman J. D. and Farnsworth K. L. (2011) Runoff, erosion, and delivery to the coastal ocean. In *River Discharge to the Coastal Ocean: A Global Synthesis*. Cambridge University Press,

New York. pp. 13-61.

Pancost R. D., Baas M., van Geel B. and Damsté J. (2002) Biomarkers as proxies for plant inputs to peats: an example from a sub-boreal ombrotrophic bog. *Org. Geochem.* **33**, 675–690.

Pancost R. D. and Boot C. S. (2004) The palaeoclimatic utility of terrestrial biomarkers in marine sediments. *Mar. Chem.* **92**, 239–261.

Ponton C., West J. A., Feakins S. J. and Galy V. V. (2014) Leaf wax biomarkers in transit record river catchment composition. *Geophys. Res. Lett.* **41**, 6420–6427.

Rommerskirchen F., Plader A., Eglinton G., Chikaraishi Y. and Rullkötter J. (2006) Chemotaxonomic significance of distribution and stable carbon isotopic composition of long-chain alkanes and alkan-1-ols in C_4 grass waxes. *Org. Geochem.* **37**, 1303–1332.

Saliot A., Mejanelle L., Scribe P., Fillaux J., Pepe C., Jabaud A. and Dagaut J. (2001) Particulate organic carbon, sterols, fatty acids and pigments in the Amazon River system. *Biogeochemistry* **53**, 79–103.

Sessions A. L. (2006) Isotope-ratio detection for gas chromatography. *J. Sep. Sci.* **29**, 1946–1961.

Sinninghe Damsté J. S., Rijpistra W. I. C., and Reichart G-J. (2002) The influence of oxic degradation on the sedimentary biomarker record II. Evidence from Arabian Sea sediments. *Geochim. Cosmochim. Ac.* **66**, 2737-2754.

Spencer R. G. M., Hernes P. J., Aufdenkampe A. K., Baker A., Gulliver P., Stubbins A., Aiken G. R., Dyda R. Y., Butler K. D., Mwamba V. L., Mangangu A. M., Wabakanghanzi J. N. and Six J. (2012) An initial investigation into the organic matter biogeochemistry of the Congo River. *Geochim. Cosmochim. Ac.* **84**, 614–627.

Spencer R. G. M., Stubbins A. and Gaillardet J. (2014) Geochemistry of the Congo River, Estuary, and Plume. In *Biogeochemical Dynamics at Large River-Coastal Interfaces* (eds. Bianchi T. S., Allison M. A., and Cai W.-J.). Cambridge University Press, New York. pp. 554-583.

Still C. and Powell J. (2010) Continental-Scale Distributions of Vegetation Stable Carbon Isotope Ratios. In *Isoscapes* (eds. West J. B., Bowen G. J., Dawson T. E., and Tu K. P.). Springer, New York. pp. 179-194.

Sun M-Y. and Wakeham S. G. (1994) Molecular evidence for degradation and preservation of organic matter in the anoxic Black Sea Basin. *Geochim. Cosmochim. Ac.* **58**, 3395-3406.

892 Sun M-Y., Wakeham S. G., and Lee C. (1997) Rates and mechanisms of fatty acid degradation
893 in oxic and anoxic coastal marine sediments of Long Island Sound, New York, USA.
894 *Geochim. Cosmochim. Ac.* **61**, 341-355.
895 Tao S., Eglinton T. I., Montluçon D. B., McIntyre C. and Zhao M. (2015) Pre-aged soil organic
896 carbon as a major component of the Yellow River suspended load: Regional significance and
897 global relevance. *Earth Planet. Sc. Lett.* **414**, 77–86.
898 van Dongen B. E., Zencak Z. and Gustafsson Ö. (2008) Differential transport and degradation of
899 bulk organic carbon and specific terrestrial biomarkers in the surface waters of a sub-arctic
900 brackish bay mixing zone. *Mar. Chem.* **112**, 203–214.
901 Vogts A., Moossen H., Rommerskirchen F. and Rullkötter J. (2009) Distribution patterns and
902 stable carbon isotopic composition of alkanes and alkan-1-ols from plant waxes of African
903 rain forest and savanna C_3 species. *Org. Geochem.* **40**, 1037–1054.
904 Volkman J., Barrett S. M., Blackburn S. I., Mansour M. P., Sikes E. L. and Gelin F. (1998)
905 Microalgal biomarkers: A review of recent research developments. *Org. Geochim.* **29**, 1163-
906 1179.
907 Volkman J., Barrett S. M. and Blackburn S. I., Eustigmatophyte microalgae are potential sources
908 of C_{29} sterols, C_{22}--C_{28} *n*-alcohols and C_{28}--C_{32} *n*-alkyl diols in freshwater environments.
909 *Org. Geochim*. **30**, 307-318.
910 Wang Y. V., Larsen T., Leduc G., Andersen N., Blanz T. and Schneider R. R. (2013a) What
911 does leaf wax δD from a mixed C_3/C_4 vegetation region tell us? *Geochim. Cosmochim. Ac.*
912 **111**, 128–139.
913 Wang Z. A., Spencer R. G., Bienvenu D. J., Mann P. J., Hoering K. A., Poulsen J. R. and
914 Holmes R. M. (2013b) Inorganic carbon speciation and fluxes in the Congo River. *Geophys.*
915 *Res. Lett.*, 511–516.
916 Wang Z. and Liu W. (2012) Carbon chain length distribution in *n*-alkyl lipids: A process for
917 evaluating source inputs to Lake Qinghai. *Org. Geochem.* **50**, 36–43.
918 Whiteside J. H., Eglinton T. I., Olsen P. E., Cornet B., McDonald N. G. and Huber P. (2011)
919 Pangean great lake paleoecology on the cusp of the end-Triassic extinction. *Palaeogeogr.*
920 *Palaeocl.* **301**, 1–17.
921 Xu, Y., Simoneit B. R. T. and Jaffé, R. (2007) Occurrence of long-chain *n*-alkenols, diols, keto-
922 ols and *sec*-alkanols in a sediment core from a hypereutrophic, freshwater lake. *Org.*

Geochem. **38**, 870-883.

925 **Table Captions**

926

927 Table 1: Weighted least squares regression correlation values (r) and significance p-values
928 between all measured C_{23+} *n*-alkyl lipid concentrations. Statistically significant (p-value ≤ 0.05)
929 correlations are bolded.

930

931 Table 2: Weighted least squares regression correlation values (r) and significance p-values
932 between all measured C_{25+} *n*-alkyl lipid $\delta^{13}C$ values. Statistically significant (p-value ≤ 0.05)
933 correlations are bolded.

934

935 Table 3: Weighted least squares regression correlation values (r) and significance p-values
936 between all measured C_{23+} *n*-alkyl lipid concentrations vs. C_{25+} $\delta^{13}C$ values. Statistically
937 significant (p-value ≤ 0.05) correlations are bolded.

938

939 Table 4: Summary statistics of plant-wax *n*-alkane, *n*-alcohol, and *n*-alkanoic acid (ΣLC_{25-35},
940 ΣLC_{26-36}) concentration and CPI data from African plant leaves (μg/g dry leaf weight).

941

942

943 **Figure Captions**

944

945 Figure 1: Map of the Congo catchment upstream of sampling location showing (a.) land cover
946 according to European Commission Joint Research Centre (Mayaux et al., 2000) and (b.) $\%C_3$
947 vs. $\%C_4$ vegetation (Still and Powell, 2010). Our sampling location is marked as a red circle. For
948 reference, Bangui Station (Coynel et al., 2005; Bouillon et al., 2012; Bouillon et al., 2014) is
949 marked as a white diamond.

950

951 Figure 2: Time series plots of (a.) Congo River discharge (Q_w), (b.) monthly fractional
952 contribution by southern hemisphere tributaries (f_{south}) as estimated by Bricquet (1993)
953 (repeating for multiple years), (c.) TSS concentration, and (d.) POC concentration measured at
954 Brazzaville/Kinshasa during the sampling period. Southern hemisphere dominated periods are
955 defined when f_{south} is greater than the median value of 39%, and are indicated by gray boxes.

Figure 3: Violin plots of *n*-alkane, *n*-alcohol, and *n*-alkanoic acid (a. – c.) POC-normalized concentrations and (d. – f.) $\delta^{13}C$ values for individual long-chain homologues. Violin plots represent the temporal distribution of values throughout the time-series using a Gaussian kernel density function. Mean values are marked as black circles, and median values are marked as horizontal black lines.

Figure 4: Time series plots of *n*-alkane concentrations – (a.) C_{29}, (b.) C_{35}, (c.) $\Sigma LC_{25\text{-}35}$ – and $\delta^{13}C$ values – (d.) C_{29} (solid line) and C_{35} (dashed line). Selected homologues are chosen to represent the increasing influence by C_4 grasses with increasing chain length. Dark gray shading represents ±1σ uncertainty, and light gray shading represents 95% confidence interval (CI). Periods when $f_{south} > 39\%$ are indicated by gray boxes.

Figure 5: Correlations between $\Sigma LC_{25\text{-}35}/\Sigma LC_{26\text{-}36}$ concentrations, ACL, and CPI vs. Congo River discharge (Q_w) measured at Brazzaville/Kinshasa for *n*-alkanes (a. – c.), *n*-alcohols (d. – f.), and *n*-alkanoic acids (g. – i.). Error bars on individual points represent ±1σ uncertainty. Black line is the WLS best-fit line, dark gray shading represents ±1σ regression uncertainty, and light gray shading represents the 95% CI. Samples collected when $f_{south} > 39\%$ are plotted as white squares, and samples collected when $f_{south} \leq 39\%$ are plotted as black circles.

Figure 6: Time series plots of *n*-alcohol concentrations – (a.) C_{24}, (b.) C_{28}, (c.) $\Sigma LC_{26\text{-}36}$ – and $\delta^{13}C$ values – (d.) C_{24} (solid line) and C_{28} (dashed line). Selected homologues are chosen to represent the autochthonous contribution to C_{24} and C_3 plant dominance of longer homologues. Dark gray shading represents ±1σ uncertainty, and light gray shading represents 95% CI. Periods when $f_{south} > 39\%$ are indicated by gray boxes.

Figure 7: Time series plots of *n*-alkanoic acid concentrations – (a.) C_{28}, (b.) C_{34}, (c.) $\Sigma LC_{26\text{-}36}$ – and $\delta^{13}C$ values – (d.) C_{28} (solid line) and C_{34} (dashed line). Selected homologues are chosen to represent the similar C_3-like isotopic composition across all long-chain homologues. Dark gray shading represents ±1σ regression uncertainty, and light gray shading represents 95% CI. Periods when $f_{south} > 39\%$ are indicated by gray boxes.

Figure 8: Correlations between $\delta^{13}C$ values vs. Congo River discharge (Q_w) measured at Brazzaville/Kinshasa for (a.) C_{24} *n*-alcohol, (b.) C_{26} *n*-alcohol, (c.) C_{24} *n*-alkanoic acid, and (d.) C_{28} *n*-alkanoic acid. Error bars on individual points represent $\pm 1\sigma$ uncertainty. Black line is the WLS best-fit line, dark gray shading represents $\pm 1\sigma$ regression uncertainty, and light gray shading represents the 95% CI. Samples collected when $f_{south} > 39\%$ are plotted as white squares, and samples collected when $f_{south} \leq 39\%$ are plotted as black circles.

Figure 9: Time series plots of the fractional contribution to the plant-wax *n*-alkyl lipid total by (a.) $\Sigma LC_{25\text{-}35}$ *n*-alkanes, (b.) $\Sigma LC_{26\text{-}36}$ *n*-alcohols, and (c.) $\Sigma LC_{26\text{-}36}$ *n*-alkanoic acids. Dark gray shading represents $\pm 1\sigma$ regression uncertainty, and light gray shading represents 95% CI. Periods when $f_{south} > 39\%$ are indicated by gray boxes.

Figure 10: Fractional contribution by (a.) $\Sigma LC_{25\text{-}35}$ *n*-alkanes, (b.) $\Sigma LC_{36\text{-}36}$ *n*-alcohols, and (c.) $\Sigma LC_{26\text{-}36}$ *n*-alkanoic acids plotted vs. Congo River discharge (Q_w) measured at Brazzaville/Kinshasa. Black line is the quadratic WLS regression line. R^2 values and significance p-values for linear (β_1) and quadratic (β_2) parameters are reported for each regression. Samples collected when $f_{south} > 39\%$ are plotted as white squares, and samples collected when $f_{south} \leq 39\%$ are plotted as black circles. Uncertainty ($\pm 1\sigma$) is smaller than the symbols for all data points.

Supplementary Table Captions:

Table EA1: Sample collection dates, TSS concentrations, %OC, and discharge measurements.

Table EA2: Concentrations of individual *n*-alkane homologues, $\Sigma LC_{25\text{-}35}$, ACL, CPI, and P_{aq}.

Table EA3: Concentrations of individual *n*-alcohol homologues, $\Sigma LC_{26\text{-}36}$, ACL, and CPI.

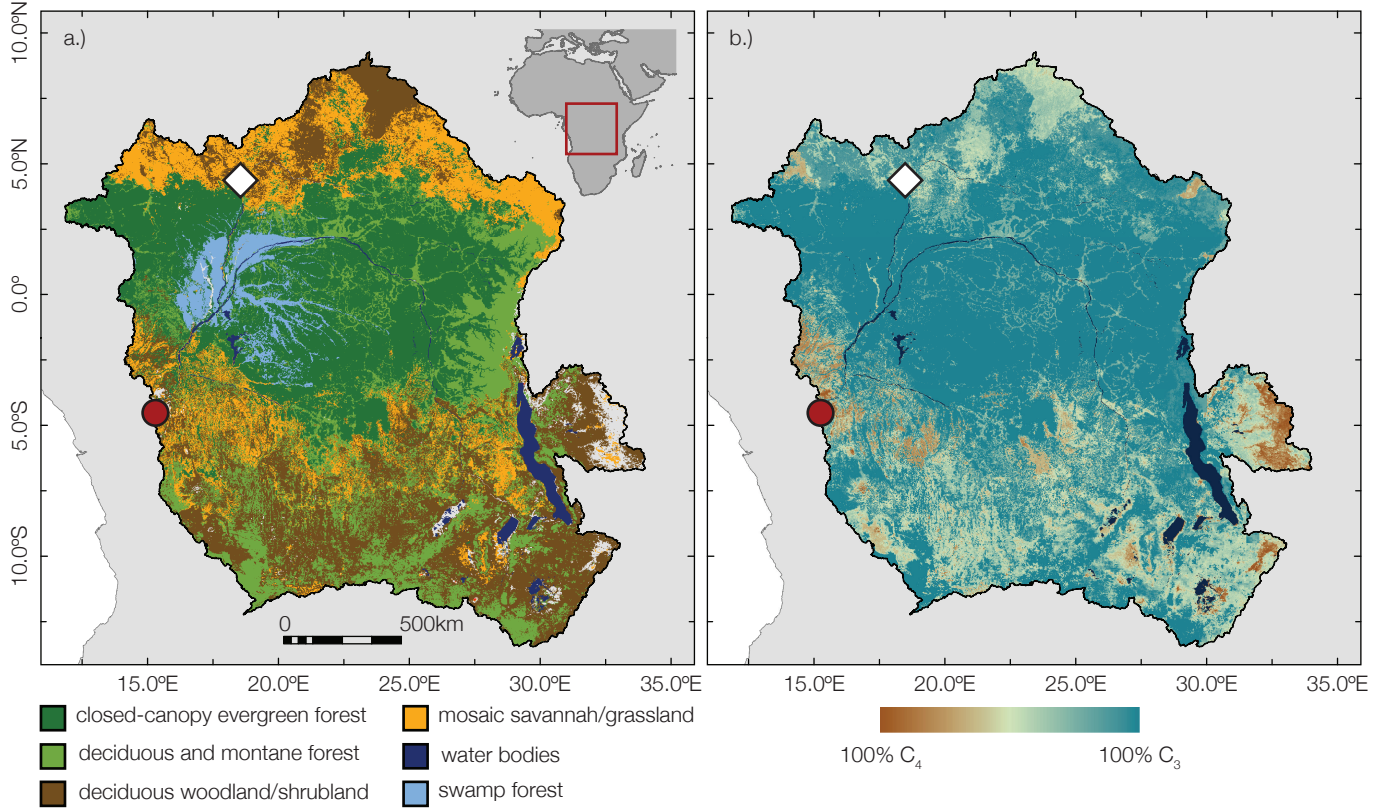
Table EA4: Concentrations of individual *n*-alkanoic acid homologues, $\Sigma LC_{26\text{-}36}$, ACL, and CPI.

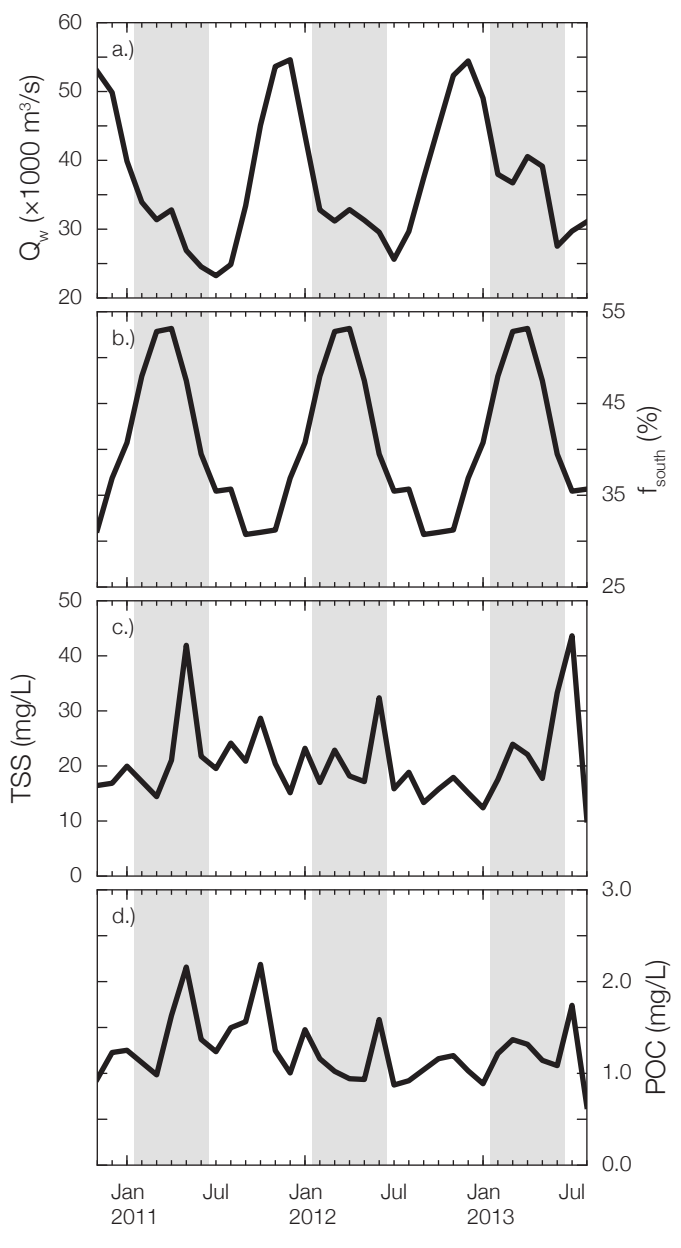
Table EA5: $\delta^{13}C$ values of individual *n*-alkane homologues.

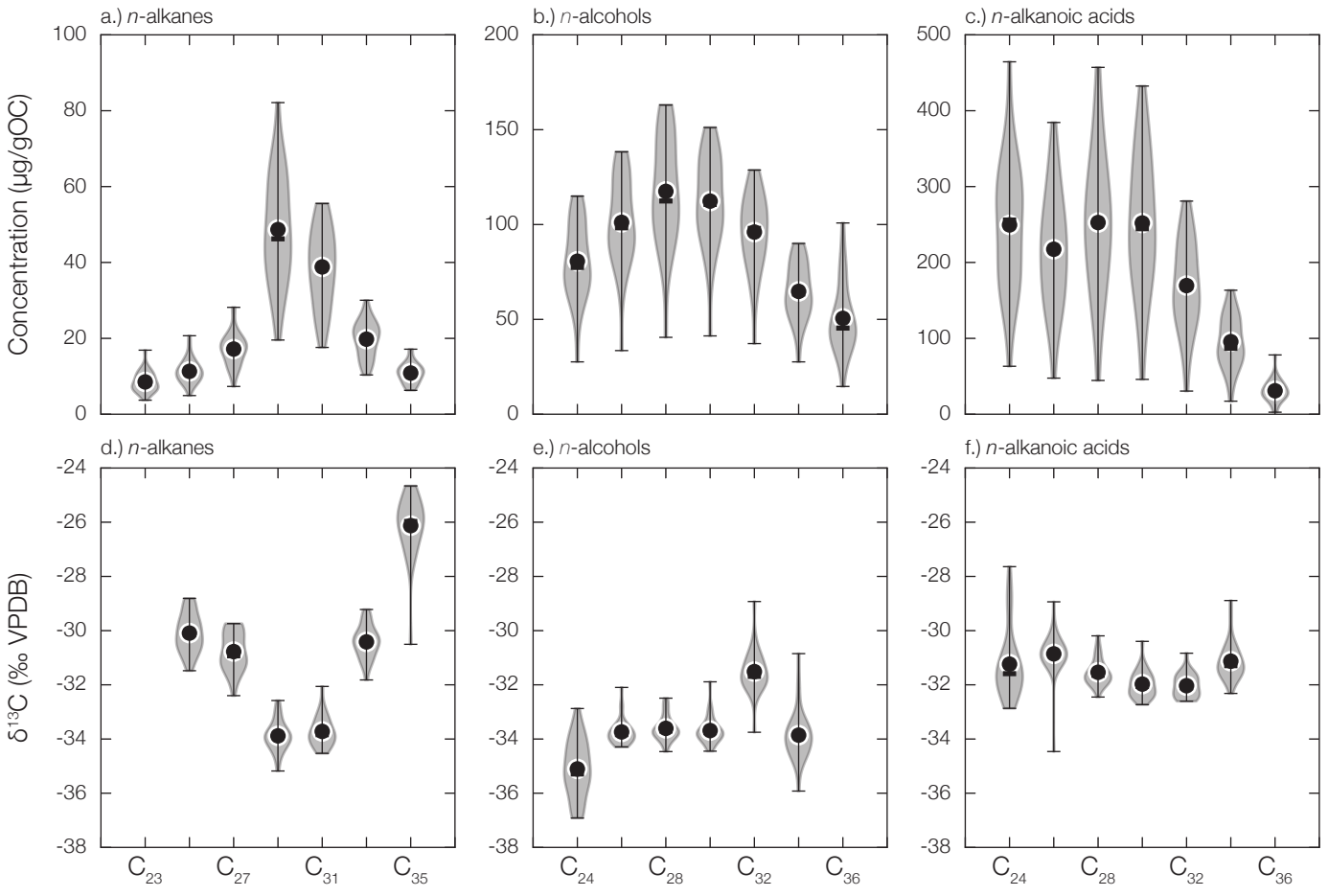
Table EA6: $\delta^{13}C$ values of individual *n*-alcohol homologues.

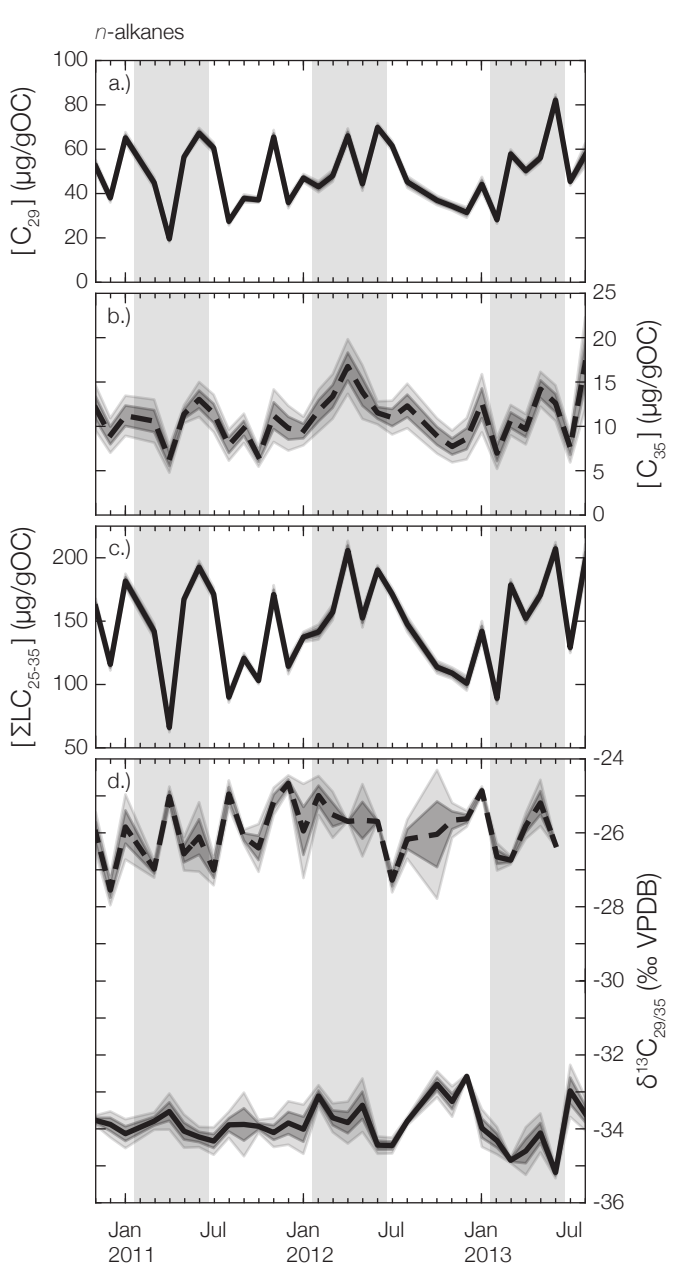
Table EA7: $\delta^{13}C$ values of individual *n*-alkanoic acid homologues.

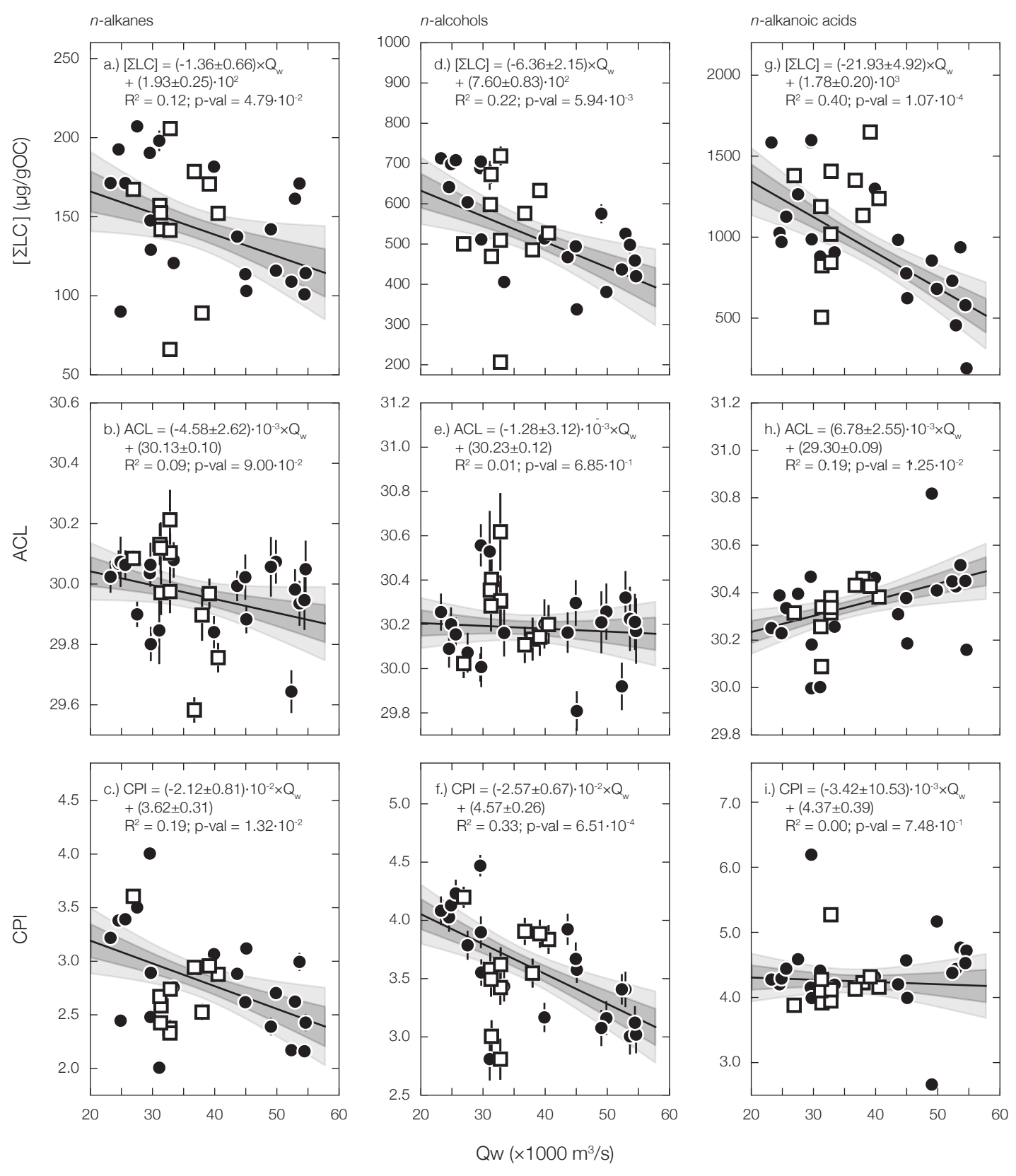
Table EA8: Fractional long-chain contribution of each compound class to the long-chain *n*-alkyl lipid total.


a.)
b.)
10.0ºN
5.0ºN
0.0º
5.0ºS
10.0ºS
15.0ºE
20.0ºE
25.0ºE
30.0ºE
35.0ºE
0
500km
closed-canopy evergreen forest
deciduous and montane forest
deciduous woodland/shrubland
mosaic savannah/grassland
water bodies
swamp forest
100% C_4
100% C_3


a.)
b.)
c.)
d.)
Q_w (×1000 m^3/s)
f_{south} (%)
TSS (mg/L)
POC (mg/L)
60
50
40
30
20
55
45
35
25
50
40
30
20
10
0
3.0
2.0
1.0
0.0
Jan
2011
Jul
Jan
2012
Jul
Jan
2013
Jul

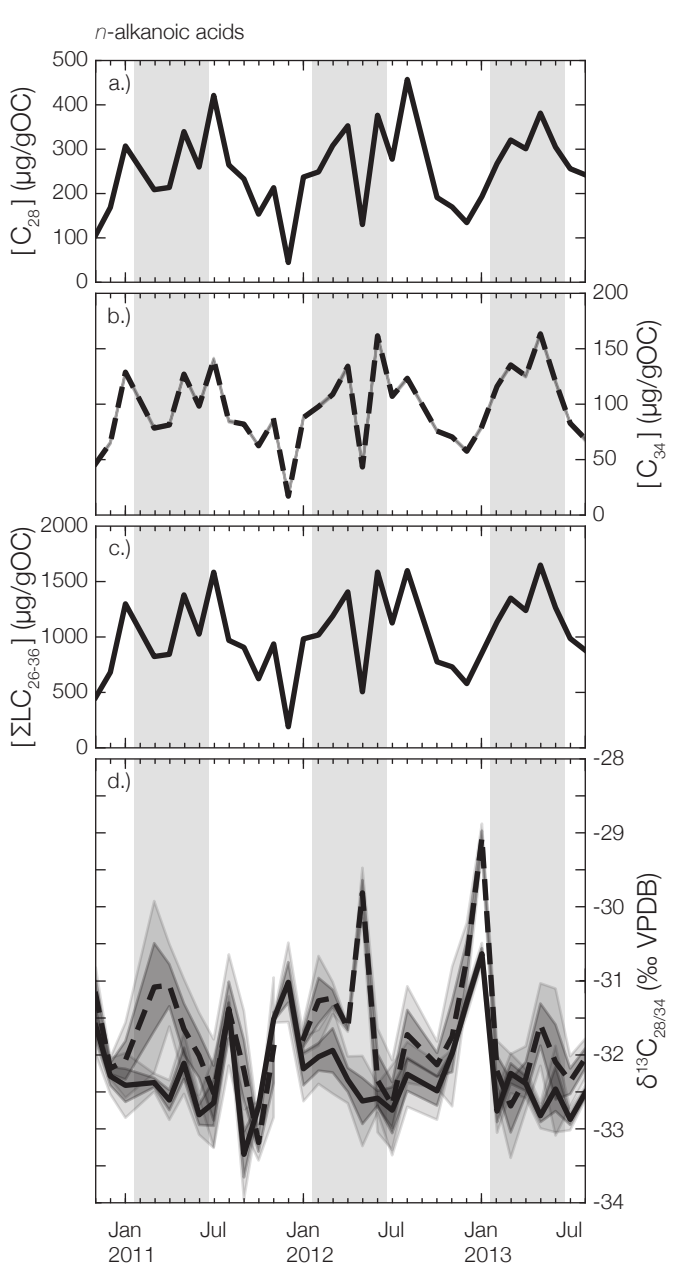

a.) n-alkanes
b.) n-alcohols
c.) n-alkanoic acids
d.) n-alkanes
e.) n-alcohols
f.) n-alkanoic acids
Concentration (µg/gOC)
δ13C (‰ VPDB)
C23
C27
C31
C35
C24
C28
C32
C36

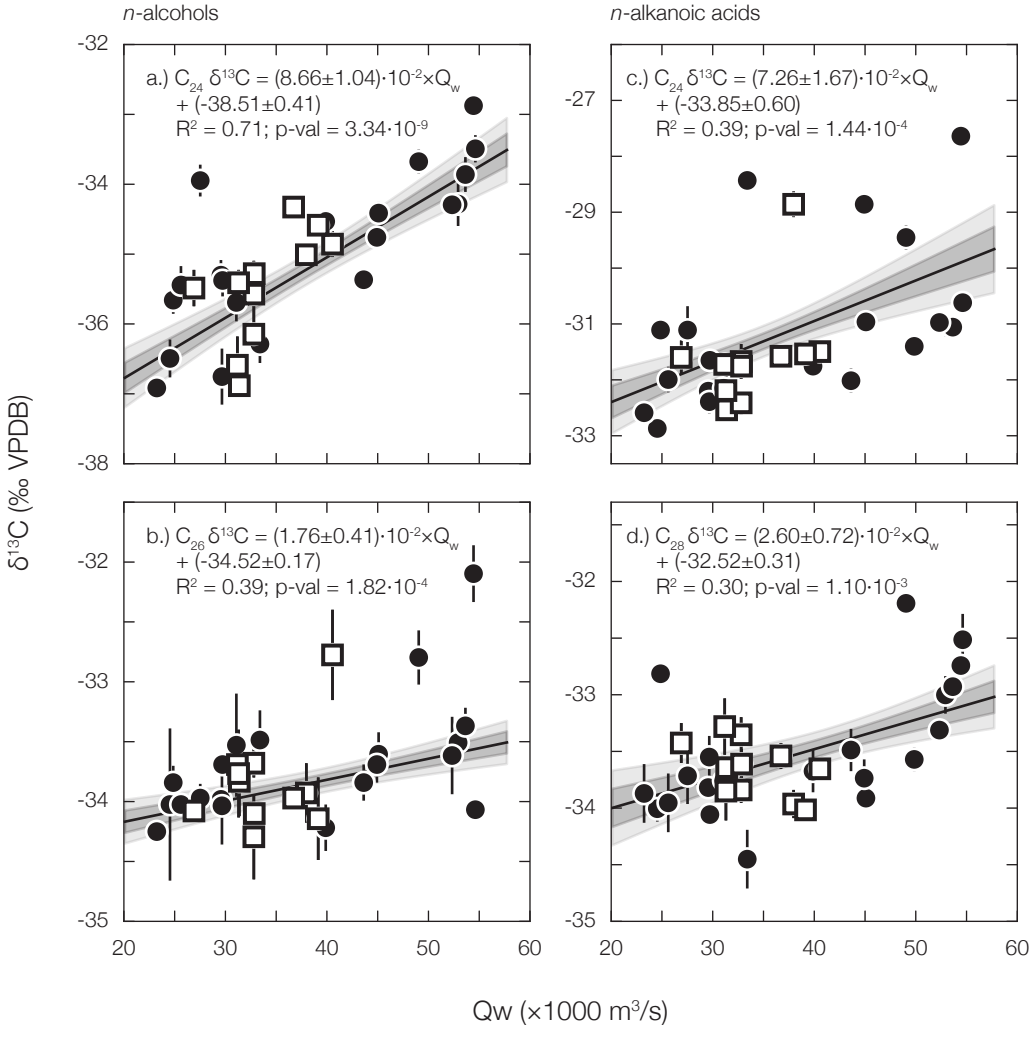

n-alkanes
a.)
b.)
c.)
d.)
[C29] (µg/gOC)
[C35] (µg/gOC)
[ΣLC25-35] (µg/gOC)
δ13C29/35 (‰ VPDB)
100
80
60
40
20
0
25
20
15
10
5
0
200
150
100
50
-24
-26
-28
-30
-32
-34
-36
Jan 2011
Jul
Jan 2012
Jul
Jan 2013
Jul

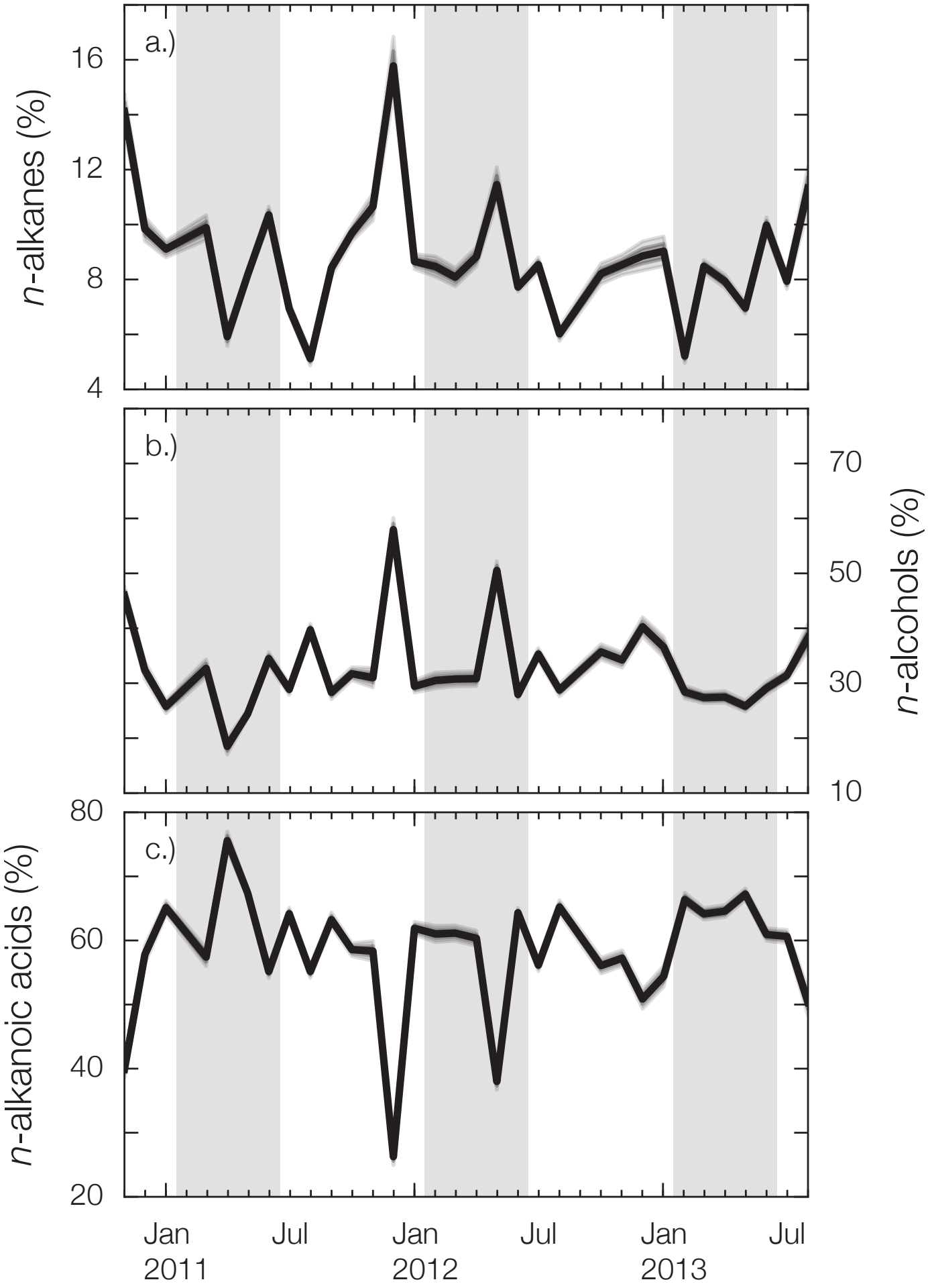

n-alkanes
n-alcohols
n-alkanoic acids
[ΣLC] (µg/gOC)
ACL
CPI
Qw (×1000 m³/s)
a.) [ΣLC] = (-1.36±0.66)×Qw + (1.93±0.25)·10²
R² = 0.12; p-val = 4.79·10⁻²
d.) [ΣLC] = (-6.36±2.15)×Qw + (7.60±0.83)·10²
R² = 0.22; p-val = 5.94·10⁻³
g.) [ΣLC] = (-21.93±4.92)×Qw + (1.78±0.20)·10³
R² = 0.40; p-val = 1.07·10⁻⁴
b.) ACL = (-4.58±2.62)·10⁻³×Qw + (30.13±0.10)
R² = 0.09; p-val = 9.00·10⁻²
e.) ACL = (-1.28±3.12)·10⁻³×Qw + (30.23±0.12)
R² = 0.01; p-val = 6.85·10⁻¹
h.) ACL = (6.78±2.55)·10⁻³×Qw + (29.30±0.09)
R² = 0.19; p-val = 1.25·10⁻²
c.) CPI = (-2.12±0.81)·10⁻²×Qw + (3.62±0.31)
R² = 0.19; p-val = 1.32·10⁻²
f.) CPI = (-2.57±0.67)·10⁻²×Qw + (4.57±0.26)
R² = 0.33; p-val = 6.51·10⁻⁴
i.) CPI = (-3.42±10.53)·10⁻³×Qw + (4.37±0.39)
R² = 0.00; p-val = 7.48·10⁻¹

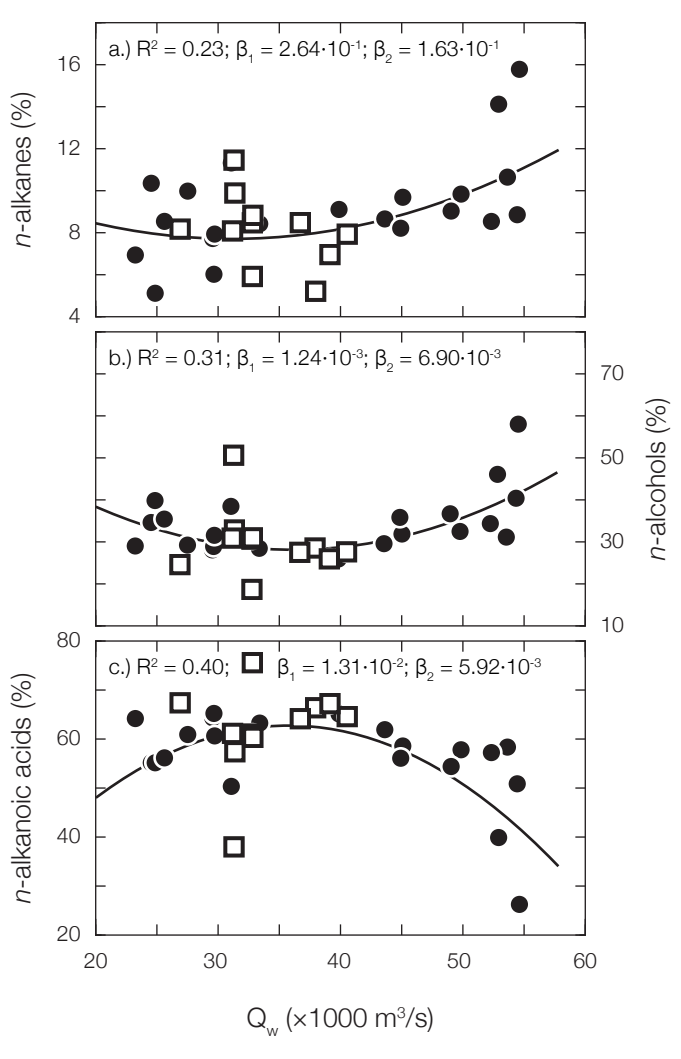
n-alcohols

a.)
b.)
c.)
d.)
[C24] (µg/gOC)
[C28] (µg/gOC)
[ΣLC26-36] (µg/gOC)
δ13C24/28 (‰ VPDB)
120
80
40
0
200
150
100
50
0
700
500
300
100
-31
-32
-33
-34
-35
-36
-37
-38
Jan 2011
Jul
Jan 2012
Jul
Jan 2013
Jul


n-alkanoic acids
a.)
[C28] (µg/gOC)
500
400
300
200
100
0
b.)
[C34] (µg/gOC)
200
150
100
50
0
c.)
[ΣLC26-36] (µg/gOC)
2000
1500
1000
500
0
d.)
δ13C28/34 (‰ VPDB)
-28
-29
-30
-31
-32
-33
-34
Jan 2011
Jul
Jan 2012
Jul
Jan 2013
Jul


n-alcohols
n-alkanoic acids
a.) C24 δ13C = (8.66±1.04)·10-2×Qw + (-38.51±0.41)
R2 = 0.71; p-val = 3.34·10-9
c.) C24 δ13C = (7.26±1.67)·10-2×Qw + (-33.85±0.60)
R2 = 0.39; p-val = 1.44·10-4
b.) C26 δ13C = (1.76±0.41)·10-2×Qw + (-34.52±0.17)
R2 = 0.39; p-val = 1.82·10-4
d.) C28 δ13C = (2.60±0.72)·10-2×Qw + (-32.52±0.31)
R2 = 0.30; p-val = 1.10·10-3
δ13C (‰ VPDB)
Qw (×1000 m3/s)


a.)
b.)
c.)
n-alkanes (%)
n-alcohols (%)
n-alkanoic acids (%)
16
12
8
4
70
50
30
10
80
60
40
20
Jan
2011
Jul
Jan
2012
Jul
Jan
2013
Jul


a.) R^2 = 0.23; $\beta_1 = 2.64 \cdot 10^{-1}$; $\beta_2 = 1.63 \cdot 10^{-1}$
b.) R^2 = 0.31; $\beta_1 = 1.24 \cdot 10^{-3}$; $\beta_2 = 6.90 \cdot 10^{-3}$
c.) R^2 = 0.40; $\beta_1 = 1.31 \cdot 10^{-2}$; $\beta_2 = 5.92 \cdot 10^{-3}$
n-alkanes (%)
n-alcohols (%)
n-alkanoic acids (%)
Q_w (×1000 m^3/s)